DOCUMENT RESUME

ED 040 514 24c EA 002 948

AUTHOR Fortune, Jim C.; And Others
TITLE Mid-South Undergraduate Research Training Program. Final Report.
INSTITUTION Memphis State Univ., Tenn.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-6-1869; BR-6-1870
PUB DATE Jul 68
GRANT OEG-2-6-061869-1355; OEG-2-6-061870-1355
NOTE 140p.

EDRS PRICE EDRS Price MF-$0.75 HC Not Available from EDRS.
DESCRIPTORS Curriculum, Educational Practice, *Educational Programs, Educational Testing, Instructional Design, Reading Instruction, *Research Methodology, *Research Skills, *Research Utilization, Special Education, Statistics, Student Teaching, Teacher Education, *Undergraduate Study

ABSTRACT

The objectives of the program were to develop personnel capable of (1) upgrading existing educational curricula through research, (2) seeking better educational practices through empirical knowledge in the use of research techniques to improve classroom instruction, (3) applying research methods to provide insights into sociological problems related to educational practice, and (4) utilizing research and evaluation techniques to improve educational practice. The program consisted of four three-credit courses, a practicum seminar in which students worked on projects, and a required course in a specialty area such as teacher training, reading instruction, or special education. GPA, progress, and postprogram placement data are provided on the students enrolled in the program. Course information includes copies of tests administered and reading lists. An appendix contains a report of two team projects conducted by the class, one dealing with the effect of the particular week day and test administration, and the other dealing with test development. [Not available in hard copy due to marginal legibility of the original document.] (Author/MLF)

BR-6-1869
PA 24
OE/BR

ED040514

Best Copy Available

FINAL REPORT

Project No. 6-1869
Grant No. OEG 2-6-061869-1355
and
Project No. 6-1870
Grant No. OEG 2-6-061870-1355

MID-SOUTH UNDERGRADUATE RESEARCH TRAINING PROGRAM

Dr. Jim C. Fortune, Director
and Instructor
Dr. Donald Thomsen, Instructor
Dr. Jack Miller, Instructor

MEMPHIS STATE UNIVERSITY
MEMPHIS, TENNESSEE

July, 1968

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

EA 002 948

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TABLE OF CONTENTS

Item . . . . . Page

SUMMARY 2

  Background 2

  Program Description 2

  Program Products 3

  Projections 3

INTRODUCTION 5

  Setting 5

  Needs 5

  Rationale 5

METHODS 7

  Objectives 7

  Eligibility 7

  Curriculum 8

  Laboratory 8

RESULTS 9

  Research Trainees 9

  Program Progress 10

CONCLUSIONS 14

  Placement of Trainees 14

  Program Future 14

APPENDICES

  A. Evaluation Materials 17

  B. Examples of Team Projects 42

SUMMARY

Background

Memphis State University is a growing state institution located in the center of the rapidly changing Mid-South area. Since its beginning as a normal school, the university has offered leadership in education to West Tennessee, Eastern Arkansas, and Northern Mississippi. Because of the newly emerging industrial culture and the resulting social revolution, the present educational problems in the Mid-South are unique to the area and present crucial educational personnel needs.

In 1966-67, the Mid-South Undergraduate Research Training Program was established through a grant from the U. S. Office of Education to develop personnel capable of and interested in (1) upgrading existing educational curricula through research, (2) seeking better educational practices through empirical knowledge in the usage of research techniques by teachers to improve classroom instruction, (3) applying research methods to provide insights into sociological problems related to educational practice, and (4) utilizing research and evaluation techniques to improve and innovate educational practice. The 1967-68 program represents an extension of the original program.

Program Description

The Mid-South Undergraduate Research Training Program maintained the following general objectives:

1. to identify undergraduate research talent;
2. to stimulate interest in and to develop appreciation of research and research training;
3. to foster the study of educational research as a field of knowledge;
4. to develop research skills;
5. to establish a firm basis for a graduate program in research;
6. to attract and to involve exceptional students in conducting research studies in a field of educational specialization, such as special education, teacher training, or reading instruction.

The Mid-South Undergraduate Research Training Program depended heavily upon a type of faculty attention to individual students and training experiences which was not probable within the structure of the traditional classroom. The program developed consisted of four three semester hour course titles taught by a team of research instructors. A research practicum in which students worked on their projects was conducted in a seminar. One course was required in one of the following specialty areas: teacher training, reading instruction, or special education.

During the first semester the students were required to attend three courses: (1) Education Tests, Measurements, and Practicum I (Education 4511), (2) Introduction to Educational Statistics I (Education 4521), and (3) Educational Statistics II and Research Design (Education 4512). During the second semester the students were enrolled in (1) Education Tests, Measurements and Practicum II (Education 4522), (2) the required specialty area course for three semester hours, and (3) a two-hour weekly seminar on research application. Three instructors team-taught the four research courses and coordinated the research seminar with the practice teaching requirements of the students.

Program Products

The acceptance of the Mid-South Undergraduate Research Training Program in the Mid-South was manifested by such events as invitations to make presentations to practitioners and students. Second, feature articles on the program were published in the major local newspaper. Finally, the successful continuation of the trainees in research serves as a manifestation of interest in and appreciation of research.

The 1966-67 program began with fifteen students who possessed an average grade point of 3.11 on a 4.00 system and who possessed ACT scores above the 80th percentile. Of the fifteen students three dropped from the program, one to take a full-time testing and counseling job, and another to become a research coordinator of a Title I ESEA program; five continued their studies at Memphis State University as research assistants; four are now completing master's degrees at other universities; one entered the teaching profession; one became a research associate in the Memphis City Schools; and one entered the insurance business. Three of the five students who remained at Memphis State University are now entering master's degree programs in research training at other universities.

The 1967-68 program began with ten scholarship students who possessed an average grade point of 3.09 on a 4.00 system and three other students whose average grade point was 2.65. All thirteen students were above the 80th percentile on the ACT. All thirteen students completed the program. Two of them are becoming research associates in the local school systems; four are planning to teach; and one of the students is being drafted. Three will remain at Memphis State University to finish their undergraduate degrees and to work as research assistants. Two plan to enter master degree programs in research training, and one is currently undecided about her future.

Projections

The Mid-South Undergraduate Research Training Program, in spite of the phasing out of the funding by the U.S. Office of Education, will be continued at Memphis State University. Part of the costs of the program to the University will be defrayed through funds designated in the West Tennessee Research Consortium to support ten students. The demand of local school systems for graduates of the research training program to serve as research associates in evaluation positions for federal and local programs has prompted this decision. A grant for a continuation of this program has been proposed as a component of Memphis State University's "Delta Flow Project" submitted under Parts C and D of the Education Professions Development Act.

INTRODUCTION

Setting

Memphis State University has in the past decade moved from a state teachers college of 4,000 students to a young, expanding university of over 15,000 students. The growth of the university is paralleled by the rapid growth of the city of Memphis and by the rapid increase of interest in education in the Mid-South. The Mid-South area is undergoing crucial changes created by urbanization, the social revolution, and modern technology.

Needs

The social, industrial, and educational changes occuring in the Mid-South have developed a comparatively virgin setting for research. The critical nature of the local educational problems, created by a changing society interacting with the unique cultural history of the area, has enticed Mid-South educators to look toward educational research for assistance. Unfortunately, there is a current shortage of research opportunities. The university research talent in the past that clashed with the conservative populace formerly migrated to settings where research was a rewarded endeavor.

Educational research activities were extremely limited until programs funded by the Elementary and Secondary Education Act, as well as other federal programs, offered supplementary financing for programs designed to meet local needs. Problems in the assessment of needs, program evaluation, and research methodology are critical needs in school systems. Until the last three years Memphis State University was offering only limited help because of its own shortage of personnel. In recent years, the University has made major strides in expanding and developing its professional personnel needs.

Rationale

Limited funds and programs for research and research training created the virtual absence of even research technicians in the Mid-South area thus making productive research difficult and too limited to be effective. An expanded college curriculum was needed giving more attention to courses in tests and measurements, statistics, and research methods. Additional research personnel were needed at the University to train personnel to operate as technicians in research projects to precede the development of large scale, meaningful research projects.

When Memphis State University began to enter into a major development of research services in 1965, the decision was made to start building at the undergraduate level and continue the development of a complete research training program through the doctorate. This program, in a comparatively virgin setting like the Mid-South, can accomplish the objective of providing the region with research technicians and highly skilled research specialists. While research specialists are required, it should also be recognized that personnel who complete only segments of the program also provide service to the area. They usually remain in the area to provide the technical personnel needed to carry on productive research. With the existence in the local school systems of skilled and semi-skilled research personnel, and with the attraction of doctorates in research to the University through the activities of research training programs, the necessary team members for productive research are now beginning to work in the Mid-South area.

METHODS

Objectives

Basically, the Mid-South Undergraduate program sought to take exceptionally talented students, offer them an accelerated instructional program in research fundamentals, and direct them toward further research training or place them in research activities in the local area. Hence, the program maintained the following general objectives:

1. to identify undergraduate research talent;
2. to stimulate interest in, and to develop an appreciation of, research and research training;
3. to foster the study of educational research as a discipline;
4. to develop research skills;
5. to establish a firm basis for a graduate program of research; and
6. to attract and to involve exceptional students in conducting research studies in a field of educational specialization, such as special education, teacher training, or reading instruction.

Eligibility

Five criteria were utilized in the selection of scholarship students:

1. each candidate was to be either a junior or senior undergraduate student capable of meeting the entrance requirements of Memphis State University (seniors were given preference in recruiting and twelve of the thirteen students were classified as seniors):
2. each candidates was to be enrolled in the full-time pursuit of a degree and was to be eligible for admittance to the teacher education program;
3. each candidate was to possess either a 3.00 grade point average on a 4.00 system or 80 percentile on the ACT test;
4. each candidate was to express a sincere interest in educational research during an interview with the program staff; and
5. each candidate was to demonstrate capability in mathematics.

Generally, the 1967-68 students were recruited by the 1966-67 students; however, there was some response to the printed brochures which were distributed.

Curriculum

The nature and history of the area dictate the importance of initial and noticeable success in research development. Through the selection of exceptional students and through the structuring of the program to provide for frequent and individual interaction of faculty and trainees, the probability of program success was increased.

Since there were no previously existing undergraduate course structures, heavy practicum applications were easy to build into the curriculum. The program consisted of (1) four specific courses in research training of which offered three semester hours credit; (2) a general seminar of practicum activities in research; and (3) an additional required course in a selected specialty area (either teacher training, reading instruction, or special education). During the first semester the students were required to complete three courses, namely; Educational Tests, Measurements, and Practicum (Education 4511); Introduction to Educational Statistics I (Education 4521); and Educational Statistics and Research Design (Education 4512).

The three first semester courses were team taught in a three hour block that met three times each week. The courses were designed to teach research and testing fundamentals and to establish a research philosophy based upon hypothesis building and critical thinking. The texts for these courses were (1) Foundations of Behavioral Research by Kerlinger, (2) Introduction to Educational Statistics by Edwards, and (3) Introduction to Tests and Measurement by Downie. The texts used in the development of fundamentals became only part of the reference library made available to the students through the Bureau of Educational Research and Services. Gage's Handbook of Research in Teaching provided a major reference source.

During the second semester the students were enrolled in Education Tests, Measurements (Education 4522); in Practicum II, the required specialty area course for three semester hours; and a two-hour weekly seminar in research application coordinated with student teaching experiences during which the students in specialty area teams pursued team projects coordinated by the seminar.

Laboratory

In order to expose the trainees to the real world of education, several laboratory experiences were established for them through the Bureau of Educational Research and Services. In addition to the experiences provided by the Bureau, the trainees participated in projects involving the Central Midwestern Regional Educational Laboratory, the Mid-South Teacher Corps, the West Tennessee Research Development Consortium, the Millington Naval Air Training Station, and Mid-South Educational Advisory Board. The students had limited participation in a small research contract on "Self-Supervision," in an in-service training program in Henry County, and in the development of a Title III ESEA proposal for a school in the Shelby County school district (Winchester).

RESULTS

Research Trainees

The 1966-67 program began with fifteen students who possessed an average grade point of 3.11 on a 4.00 system. Each student who had an ACT score, scored above the 80th percentile. Of these fifteen students, two dropped the program to accept full-time employment in research activities and one dropped because of the illness of his wife. Of the twelve students who completed the program, five remained at the University as research assistants to complete their undergraduate degrees. Of the four program graduates who entered graduate study in other universities, two gained scholarships and two received assistantships. Of the reamining trainees, one entered the teaching profession, one was hired as a research associate in the research division of the Memphis City School System, and one entered the insurance business.

Candidate Data 1966-67

Names	GPA 4 pt. system	ACT Composite	Status
Selections:			
1. Truman L. Atkins	3.08	95.0	Senior
2. Helen P. Best	2.94	81.0	Senior
3. Virginia W. Blanton	2.96	83.0	Senior
4. Terry Ann Bond	2.98	99.0	Junior
5. Phyllis A. Burdette	3.14	No record	Senior
6. Kathy Marie Eggers	2.91	99.0	Senior
7. Pallie B. Hamilton	3.41	No record	Senior
8. Roy A. Hardy	3.28	97.0	Senior
9. Lawrence N. Koss	2.99	84.0	Junior
10. Nodya Jean Loftiss	2.72	94.0	Junior
*11. Shirley McClintock	3.56	No record	Junior
12. Patricia A. Teague	3.75	No record	Senior
**13. Dana Thrasher	3.48	No record	Senior
14. Mary Lou Wallace	2.82	80.0	Senior
15. Thomas Carl Watson	3.02	No record	Senior
Alternates:			
***16. Thomas Ray Greer	3.04	No record	Senior
***17. Patricia Seymore	2.62	94.0	Junior
*18. Almeda M. Zent	3.00	78.0	Senior

*Had to drop program because of scheduling problem.

**Did not accept program because of unseen school program commitments (activity honors).

***Scholarships were transferred to these two alternates.

The 1967-68 program began with ten scholarship students who possessed an average grade point of 3.09 on a 4.00 system and three other students whose average grade was 2.65. All thirteen students possessed a minimum ACT percentile of 80 or above and all completed the program. Two of them have been employed as research associates in local school systems. Four plan to teach, and one is being drafted. Three will remain at Memphis State University as undergraduate research assistants. Two plan to do graduate study in research training, and one is undecided about his future.

Candidate Data 1967-68

Names	GPA 4 pt. system	ACT Composite	Status
Selections:			
1. Theresa A. Browning	3.00	92.0	Senior
2. Lee Elizabeth Carothers	3.20	89.0	Senior
3. Patricia L. Couch	3.18	90.0	Senior
4. Mary Jo Dye	3.83	95.0	Senior
5. Janice Gale Goff	2.99	91.0	Junior
6. Charles H. Holland, Jr.	2.76	84.0	Senior
7. Lea W. Joyner	2.83	85.0	Senior
8. Patricia A. Mahaffey	3.06	90.0	Senior
9. Charlene M. Mott	3.04	86.0	Senior
10. Lucy T. Reap	3.00	92.0	Senior
Other Students:			
11. Louida McCombs	2.69	82.0	Senior
12. Walter Mathis	2.62	80.0	Senior
13. Lynn Wells	2.65	85.0	Senior

Program Progress

In order to evaluate the program an adjective check list and an open-ended questionnaire were given the participants. The results of these two data sources with other data sources, such as results of content testing, will be reviewed in regard to program objectives.

The undergraduate research training program was able to identify undergraduate research talent. The excellence of grade point averages, the number of the trainees remaining in research, and the performance of the research training students on the materials and tests developed by Baker and Sullivan (mean scores were slightly above the mean scores of the participants in the AERA workshops in New York, 1967), substantiate that objective one, talent identification, was met. However, there exists some disappointment in that scholarships were required to interest the students to enter the program, even though 80 percent of them reported that they would have taken the program without a scholarship if they had had proper program expectancy.

Newspaper articles, program inquiries, and increased research activity indicate that interest in research was stimulated. The continued study of research by program participants offers evidence that an appreciation of research was developed. Unfortunately, this interest has failed to generalize to most of the area. Continued efforts can be expected to gain the objective of generalized regional interest.

Students in research training now recognize educational research as a legitimate field of knowledge. The continuation of the program and the development of additional graduate courses in educational research will substantiate this conclusion and will justify offering an opportunity to develop research skills. The undergraduate research training program has paved the way for the establishment of excellent field study laboratories, stronger graduate courses in research, and competent personnel to serve the graduate study of research.

On the adjective check list ten questions were asked. Five bi-polar adjectives were assigned to each question with a nine-point continuum between each. Twelve of the thirteen students responded to the questionnaire. The bi-polar adjectives are listed below each question, and the response pattern is shown between the adjectives.

1. The effect of the entire research training program on anticipated classroom teaching:

GOOD 9 : 2 : 1 : 0 : 0 : 0 : 0 : 0 : 0 BAD

USELESS 0 : 0 : 0 : 0 : 0 : 0 : 0 : 3 : 9 VALUABLE

POSITIVE 10 : 2 : 0 : 0 : 0 : 0 : 0 : 0 : 0 NEGATIVE

IMPRACTICAL 0 : 0 : 0 : 0 : 0 : 0 : 2 : 3 : 7 PRACTICAL

IMPORTANT 7 : 4 : 0 : 0 : 1 : 0 : 0 : 0 : 0 UNIMPORTANT

2. Influence of the program on my future career:

CHANGED 3 : 0 : 2 : 1 : 0 : 1 : 0 : 1 : 4 UNCHANGED

USELESS 0 : 0 : 0 : 0 : 0 : 1 : 1 : 1 : 9 VALUABLE

POSITIVE 5 : 3 : 3 : 1 : 0 : 0 : 0 : 0 : 0 NEGATIVE

IMPORTANT 4 : 2 : 4 : 2 : 0 : 0 : 0 : 0 : 0 UNIMPORTANT

GOOD 2 : 3 : 2 : 0 : 0 : 0 : 0 : 0 : 0 BAD

3. Class organization: time blocks, current pattern of 3 sessions first semester, 1 session second semester:

GOOD 0 : 2 : 2 : 1 : 2 : 1 : 2 : 0 : 2 BAD

IMPRACTICAL 1 : 1 : 2 : 1 : 3 : 0 : 1 : 1 : 2 PRACTICAL

CONVENIENT 1 : 1 : 0 : 1 : 3 : 2 : 0 : 2 : 2 INCONVENIENT

USELESS 0 : 0 : 0 : 1 : 2 : 4 : 1 : 3 : 0 VALUABLE

SATISFYING 1 : 2 : 0 : 2 : 3 : 2 : 1 : 1 : 0 DISAPPOINTING

4. Class assignments, homework, semester projects:

GOOD 4 : 7 : 0 : 1 : 0 : 0 : 0 : 0 : 0 BAD

PRACTICAL 6 : 5 : 1 : 0 : 0 : 0 : 0 : 0 : 0 IMPRACTICAL

USELESS 0 : 0 : 0 : 0 : 0 : 0 : 1 : 5 : 6 VALUABLE

SATISFYING 4 : 3 : 1 : 2 : 1 : 1 : 0 : 0 : 0 DISAPPOINTING

IMPORTANT 7 : 3 : 2 : 0 : 0 : 0 : 0 : 0 : 0 UNIMPORTANT

5. Text materials, library guidance, reference materials, mimeographed materials, etc.

BAD 0 : 0 : 0 : 0 : 0 : 0 : 2 : 5 : 5 GOOD

USELESS 0 : 0 : 0 : 0 : 0 : 1 : 1 : 5 : 5 VALUABLE

PRACTICAL 3 : 6 : 2 : 0 : 1 : 0 : 0 : 0 : 0 IMPRACTICAL

SATISFYING 3 : 1 : 7 : 0 : 0 : 0 : 1 : 0 : 0 DISAPPOINTING

ORGANIZED 6 : 5 : 0 : 0 : 1 : 0 : 0 : 0 : 0 DISORGANIZED

6. Class structure, professional climate, formality, personal attention to profession-related problems:

BAD 0 : 0 : 0 : 0 : 0 : 0 : 1 : 2 : 9 GOOD

USELESS 0 : 0 : 0 : 0 : 0 : 0 : 1 : 4 : 7 VALUABLE

PASSIVE 0 : 0 : 0 : 0 : 1 : 1 : 0 : 3 : 7 ACTIVE

FRIENDLY 8 : 2 : 2 : 0 : 0 : 0 : 0 : 0 : 0 UNFRIENDLY

SATISFYING 6 : 4 : 2 : 0 : 0 : 0 : 0 : 0 : 0 DISAPPOINTING

7. Recruitment literature and recruitment activities:

ACCURATE	1	2	2	0	0	2	0	1	3	MISLEADING
STIMULATING	0	1	4	1	0	1	0	2	2	DULL
UNORGANIZED	0	0	0	0	1	2	3	3	2	ORGANIZED
UNDERSOLD	3	2	1	2	3	0	0	0	0	OVERSOLD
IMPORTANT	0	5	2	0	2	0	0	0	2	UNIMPORTANT

8. Relationship to other course offerings of the college and to other professors on campus:

REPLICATION	0	0	0	0	0	0	1	0	11	UNIQUE
DULL	0	0	0	0	0	0	1	3	8	STIMULATING
IMPORTANT	7	3	2	0	0	0	0	0	0	UNIMPORTANT
GOOD	7	3	2	0	0	0	0	0	0	BAD
ACCEPTING	7	3	1	1	0	0	0	0	0	REJECTING

9. Experience in dealing with people, classroom observance, working with school pupils (kindergarten through 12), awareness of teaching conditions:

HELPFUL	11	1	0	0	0	0	0	0	0	UNHELPFUL
RIGHT	7	2	2	0	1	0	0	0	0	WRONG
IMPRACTICAL	0	0	0	0	0	1	0	2	9	PRACTICAL
VALUABLE	8	3	0	0	1	0	0	0	0	USELESS
REVEALING	10	1	0	0	1	0	0	0	0	STEREOTYPED

10. Preparation for education-related career:

COMPLEMENTARY	10	2	0	0	0	0	0	0	0	UNRELATED
EXCESSIVE	0	1	0	0	2	1	0	1	7	SUPPLEMENTARY
UNHELPFUL	0	0	0	0	0	0	0	2	10	HELPFUL
IMPORTANT	7	5	0	0	0	0	0	0	0	UNIMPORTANT
PASSIVE	0	0	0	0	0	0	1	4	7	ACTIVE

CONCLUSIONS

Placement of Trainees

With the existing void in the teaching ranks of skilled evaluators, there exists a competitive field of placement for students having research skills. The role of evaluation technician has a place in education if adequate personnel can be trained to fill it. Teacher weaknesses in testing procedures, increased evaluation demands in federal programs, and expanding alternatives in educational programming have generated critical shortages in evaluation personnel.

The undergraduate research training serves to supply educational institutions with some of these technicians. There appears also to be an adequate routing of graduates toward doctorial-level study of research theory and methdology. The research-training undergraduates have few problems in finding financial support for further study.

Program Future

Research recruitment can be successfully started at the undergraduate level. Not only are the by-products of normal program attrition beneficial to the world of education, but also an adequate percentage of students continue research training on the graduate level. The students who continue will have the fundamentals that are essential for the forwarding of research as a field of knowledge.

At Memphis State University, the undergraduate program will be retained, but on a limited scope because of the extensive costs of the program. This limited scope should only be temporary because demand will soon overcome any reticence to commit larger sums to the program. In order to meet the needs of modern education with quality, heavy investment is often demanded for adequate training. As the current trainees validate the adequacy of their training, greater investments will be forthcoming.

APPENDIX A

Evaluation Criteria

Midterm, Education 4521 - Introduction to Educational Statistics I
Final, Ecucation 4521 - Introduction to Educational Statistics I

Midterm, Education 4522 - Introduction to Educational Statistics II
Final, Education 4522 - Introduction to Educational Statistics II

Final, Education 4511 - Education Tests, Measurements and Practicum Application I
Final, Education 4512 - Education Tests, Measurements and Practicum Application II

15/16

Evaluation Criteria
Midterm Education 4521

Miss Jones was a fifth-grade science teacher in an overcrowded school. In order to develop a motivational force she allowed her class to pick a three-weeks topic and to select their own methods of learning the topic. The class selected space travel and suggested that knowledge of space travel could be gained through reading, motion picture films, programmed learning, field trips, and simulation. Miss Jones wanted to take full advantage of the student's desire to learn so she devised an experimental treatment of the unit. She randomly divided the class of thirty-five into five groups of seven each and taught each group with one of the suggested methods. In order to find differences in the treatments she gave each group the same ten-item criterion test. The following results were obtained:

Group I (Reading)		Group II (Motion Pictures)		Group III (Programmed Learning)		Group IV (Field Trips)		Group V (Simulation)	
Student	Grade	Student	Grade	Student	Grade	Student	Grade	Student	Grade
1	9	8	4	15	6	22	8	29	8
2	7	9	5	16	6	23	9	30	8
3	7	10	5	17	5	24	6	31	6
4	8	11	6	18	9	25	5	32	5
5	6	12	8	19	7	26	4	33	4
6	4	13	10	20	3	27	2	34	1
7	5	14	10	21	10	28	10	35	9

1. Describe Miss Jones; experimental design.

2. Discuss ways of improving the design using the same sample.

3. What was meant by using the experimental design to take advantage of the students' desire to learn?

4. Calculate the classes' mean and variance.

5. Calculate the sampling error variance for each group.

6. Calculate the mean for each group.

7. Calculate the between groups variance? Can this variance be classified as experimental variance?

Evaluation Criteria
Final Education 4521

I. The following is an excerpt from the study done by W. W. Charters, Jr. at Washington University and printed in the American Educational Research Journal, Vol. II, No. 3, May, 1965. The study is titled, "The Relation of Morale to Turnover to Teachers."

Studies of industrial morale conducted by the survey research center of the University of Michigan a number of years ago showed that morale, as conventionally measured, consiste of at least four factorially independent dimensions - intrinsic job satisfaction, satisfaction with the employing company, liking for the supervisor, and satisfaction with one's mobility within the company (Katz and Con, 1952). In a recent study, the author constructed brief, cumulative scales to measure the first two of the components of morale in the educational setting (Charters, 1964). One scale measured intrinsic satisfaction derived from teaching; the other, the teachers identification with the school system. While these scales showed meaningful relationships with other attitudinal and background variables in the initial study, it seems advisable to seek stronger validation of the measures by examining their capacity to predict teacher movement out of the school system. The present article reports on the occupational status of 538 classroom teachers, three years after they have been measured on the two scales.

When the follow-up study began, the investigators planned to predict differential predictions from two scales regarding the direction of teacher mobility. They expected that with a teacher that is high on identification with the system and low on satisfaction of teaching would move into administrative or other staff positions within the same system; while a teacher who is low on identification with the system but high on satisfaction with teaching would move to classroom position in other school districts. Teachers with low scores on both scales were expected to leave education altogether or at least both the system they were in and classroom teaching; and those high on the two scales would be immobile. Unfortunately, it was not possible to test the ability of the scales to discriminate among the various kinds of mobility; too few teachers have made the necessary kinds of moves. Only five teachers, for example, have moved to non-teaching jobs within the system. Consequently, the study was limited to examining the relationship of each scale separately to the simple incidents of movement out of the system during the three year period, ignoring the destination of those that had moved.

1. Ignoring Dr. Charters' disappointment with the lack of mobility in his sample describe a research design which you might envision being used in the study of this problem.

2. In the analysis of the data that Charters had collected for this study he used basically a comparison design for analysis. In his design he categorized by males, single females, and married females, the teachers in the sample. He then proceeded to test the difference between the means of those who remained in the system versus those who left the system. His analysis between mean differences was made by one tailed t-tests. Describe the appropriateness or lack of appropriateness of this statistical treatment.

II. The text utilized in the research training class, Foundations of Behavioral Research by Fred M. Kerlinger has been reviewed by several reviewers. Don Medley of the University of New York in reviewing the book states:

> The concept of the "scientific approach" put forth in the first two chapters of the book seems to be something like this: science proceeds by (1) proposing hypothesis deduced by theory (or other hypothesis); (2) testing the hypothesis; which (3) shows the hypothesis to be "probably correct" or "probably incorrect;" and (4) confirms (or refutes) the theory. The construction of theory is the ultimate goal of science to the present review; this seems to imply that the probability is interpreted as referring to subjective probabilities, that is the strength of belief that the scientist feels in his hypothesis and his theories, and that the whole idea of science is to strengthen that belief. This position is inconsistent with the point of view that least some behavior scientists which goes something like this: science proceeds by 1. proposing a hypothesis (which may be based upon theory) that is "exact" - that is, from which the results for an experiment can be predicted and 2. whose rejection will necessary imply the acceptance of some alternative to the hypothesis (usually only the hypothesis that fits these conditions is a null hypothesis); 3. conduct an experiment and reject the hypothesis if the results are inconsistent with the prediction; and if they are 4. accept the alternative hypothesis. Theory is built out of such indirectly supported hypothesis. No probability can be assigned to the truth of a hypothesis except one (if it is true or if it is false). 5. prediction of behavior; theories generally improve the ability to predict and are therefore useful as means to an end, but not in themselves.

In the review of the Kerlinger book, a reference is made to the interpretation of randomization by Kerlinger.

> Kerlinger's basic point of view leads to a variety of unacceptable positions to a reader who disagrees with his philosophy of research. For example, the role of randomization

is described as that of equalizing charazteristics not otherwise controlled; if data is submitted to a randomization process or distributed according to chance expectatinns, it is correct to say "that the principle of randomization has operated successfully."

To one who sees a purpose of randomization as validation of the estimate of error it is nonsense to talk about randomization as successful or unsuccessful. Is it not the role of the test of significance or the level of competence to allow for the ever present possibility that randomization has not been "successful?"

1. React to the disagreement of Dr. Kerlinger and Dr. Medley concerning "scientific approach."

2. Explain the differences in the two philosophies toward randomization.

III. Compute Between, Within, and Total Variance for these three groups: Is the experimental variance greater than the error variance?

I	II	III
8	1	5
7	4	4
5	6	3
4	8	2
3	9	1
9	12	8
12	1	9
11	7	7
8	6	3
2	8	2

IV. Correlate the two frequencies: How much variance of x can be explained by y?

x	y
21	20
36	30
12	15
15	18
18	16
12	9
20	19
15	14
7	10
8	10

Evaluation Criteria
Midterm, Education 4522

1. Define: independence, mutual exclusion, and exhaustiveness.

2. Suppose that you are sampling ninth-grade youngsters for some research purposes. There are 250 ninth graders in the school system, 130 boys and 120 girls.

   (a) What is the probability of selecting any youngster?
   (b) What is the probability of selecting a girl? A boy?
   (c) What is the probability of selecting either a boy or girl?
   (d) In a sample of two children what is the probability of selecting first a boy and then a girl?
   (e) Suppose you drew a sample of 100 boys and girls. You got 90 boys and 10 girls. What conclusions might you reach? Why?

3. An educational investigator drew a random sample of 64 ninth-grade students in a junior high school and asked them if they favored monthly conferences with their homeroom teacher. Of the 64 students, 40 said Yes.

   (a) Is this a "real" majority opinion? Is it a statistically significant response?
   (b) Test the response for significance. Use X^2, and assume equiprobability (that is, if the students answered Yes and No at random, what responses would you expect to get?) An X^2 as great as, or greater than, 3.84 is significant at the .05 level, 6.64 at the .01 level.

   (c) Test the response using binomial statistics and the normal probability curve. (Use Kerlinger Chapter 10).

4. $\bar{X} = 150$, $S_x = 28$, $n = 16$, test the hypothesis that $u = 135$. Compute the value of t and tell what you do with the hypothesis.

5. Suppose an experimenter believes that the number of pellets rats will eat in a 10-minute period is related to the length of time since they last ate. To test his hypothesis he randomly selects three groups of six rats each. After giving each group preliminary training he tests group 3, 3 hours after eating; group 12, 12 hours after eating; and group 24, 24 hours after eating. The results of his experiment are presented below with the data being the number of pellets eaten by each animal in a 10-minute period.

GROUP 3	GROUP 12	GROUP 24
0	5	5
7	2	11
2	3	9
1	3	9
1	6	4
7	5	10

(a) What statistical analysis is in order?
(b) Describe the assumptions required by this analysis.
(c) Analyze and interpret the data.

6. Do a double-classification ANOVA on:
Reading Scores for a Stratified Random Sample

Location	Sex	
	M	F
URBAN	4	1
	9	4
	9	5
	10	6
RURAL	3	4
	7	4
	7	8
	7	4

7. Compute the product-moment correlation between shoe size and height in inches.

SUBJECT	SHOE SIZE	HEIGHT
1	9.0	70
2	6.5	62
3	8.0	66
4	7.5	69
5	8.5	72

8. From $S^2 = \frac{1}{n}\Sigma(\bar{X}-X)^2$ derive $S^2 - \frac{1}{n}\Sigma X^2 - X$

9. Koenker, in his study of arithmetic readiness, reports, among other data, difference scores. These scores are the differences between the final gain scores of the matched pairs of his experimental and control groups on an arithmetic test. Among the 27 scores, 24 are positive and 3 are negative, that is, 24 of the experimental group S made greater gains than their control-group partners did, and 3 of them did not gain as much as their control-group partners.

(a) Use the sign test (14.1-14.3-14.4 Kerlinger) to test the significance of the result. Interpret.

(b) Koenker reports a t ratio of 5.23, significant at the .01 level - actually significant at the .001 level. How does the result of the sign test compare with the t-test?

10. Miss Jones utilized two methods of feedback on her written comments being returned with a diagnostic test. In the examination of the class, Miss Jones wished to see if Treatment A, praise centered comments, produced better grades than Treatment B, blame centered comments. The scores on the examination were:

Treatment A	Treatment B
3	6
5	5
2	7
4	8
6	9
2	4
7	7
	8
	9
	7

Random techniques were used

11. If a large population is to be studied by choosing a sample from it, are the two major requirements that must be met in selecting the sample? What kinds of useful conclusions depend for their justification on carefully satisfying these requirements?

12. A jury hearing a murder trial is working through the evidence toward a decision centered on the hypothesis: The defendant is not guilty. What two errors are possible in the decision, and which is the most serious? Justify your answer.

13. What are the reasons for the almost universal use of critical regions of sizes .05 and .01? Would any of these reasons be valid arguments against the use of regions of size .03, .07, or .20? State your reference.

14. In a study of teaching methods the following data was found:
Traditional method $\underline{X} = 75, \sigma x = 5.4, n_x = 30$
Experimental method $\bar{Y} = 79, \sigma y = 5.1, n_y = 32$

Using X = .05 as in the example, decide whether to accept or reject Ho: ux = uy.

15. A teacher with 135 students estimates on the basis of past experience that twenty students will earn an A, twenty-five a B, forty-nine a C, twenty-five a D, and sixteen an F. The grades given at the end of the term included fifteen A's, twenty-one B's, forty-two C's, thirty-one D's, and twenty-six F's.

(a) Frame suitable null and alternative hypotheses which may be tested using the X^2 test with $\alpha = .05$.
(b) Find X^2 and the number of degrees of freedom.
(c) What conclusions or decisions can be made concerning the method of estimating grades and the class to which the method was applied?

16. Plot a scattergram and compute a correlation coefficient for:

Student Number	English Grade	Vocabulary Grade
1	94	92
2	89	90
3	85	78
4	81	85
5	79	81
6	78	83
7	76	75
8	75	74
9	71	82
10	70	69
11	68	71
12	65	63
13	60	58
14	57	64
15	55	58
16	49	52
17	48	45
18	39	35

17. Find rank order r for the rankings:

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Y	4	11	3	12	2	1	13	7	14	6	5	15	10	16	9	8

18. An urn contains six red and four black balls. Two balls are drawn without replacement. What is the probability that the second ball is red if it is known that the first is red?

19. Given the samples (1.8, 2.9, 1.4, 1.1), (5.0, 8.6, 9.2) from normal populations, test whether the variances are equal at the .05 level.

20. The following data satisfy the two-way classification model

$Yij - u + d_i + e_j + e_{ij}$ $i = 1, 2, 3, 4$
$j = 1, 2, 3, 4, 5, 6$

Factor A (e)

Factor B ( )

	1	2	3	4	5	6	TOTAL
1	21	17	56	59	41	51	
2	20	19	61	62	46	55	
3.	24	23	54	54	39	50	
4	14	18	56	55	42	48	
TOTAL							

Evaluation Criteria
Final, Education 4522

1. The twentieth century is increasingly demanding people with high levels of education. In the future the lowest level of education and competnece will be above what it is today. Much educational practice in the United States assumes that only a limited proportion of American youth is capable of high level educational attainment. Such an assumption may drastically limit the production of necessary education personnel.

   The student's self-concept of his ability as a school learner may limit his learning and may prevent his achievement of a higher educational level. Theoretical support for the investigation of self-concept is found in the writings of Mead, Combs, and Snygg. Research has indicated that there is a consistent relationship between self-concept and behavior. It is hypothesized that the child performs in the manner that is consistent with his self-concept which is acquired during interaction with significant others who hold expectations of the student as a learner. (From pp. 280 - Galfo and Miller, Interpreting Education Research).

An elementary school system has volunteered to allow you to use their forty-two third grade classes (average thirty) for three years to study the relationship of self-concept to academic achievement. Describe the hypotheses, variables, procedures and analyses which you would utilize given this opportunity. Critique your design.

2. The term "team teaching" has been used in recent years to describe processes by which pupil instruction becomes a group rather than an individual effort. Unfortunately, although some research has been made on the use of team teaching, there still is not much information available concerning many aspects of the concept. Apparently the team idea has evolved from research into the general area of teacher utilization. (Galfo and Miller, 319).

   A junior high school district has contacted you to aid them in planning staff innovations centered around team teaching. They wish to establish the most appropriate formats, organizations and procedures available in their team teaching system. Unfortunately, there is little information available which might indicate the ways which team teaching can best be utilized in their district. Aid them by designing a multigroup experimental study using their sixty junior high schools and providing them with empirical data which they may use in answering questions concerned with teacher morale, student achievement, and administrative frustration.

3. Flexible scheduling offers to the student independent study time rather than study hall periods. In independent study the student is offered freedom of movement, opportunity for personal scheduling of work time, opportunity to use study carrels or resource centers, and choice of study modules. Design a descriptive study which would add to the current state of knowledge concerning independent study; stating problem, hypothesis, analysis, and anticipated roadblocks assuming an unlimited sample.

4. Design an experimental study using the same problem but restating the hypothesis, design procedures and analysis in proper form.

PLEASE: Write a candid and factual evaluation of the course and program. Include strong points, weak points, and suggested revisions. Has the program altered your career aspirations? Do you think you will seek further learning and future occupation in research? This can be turned in anonymously if desired.

Evaluation Criteria
Final, Education 4511

This problem is to be solved by two of you working together. Using the following data, tabulate the scores and group the data according to the standards suggested in class and determine the following:

A. Median score
B. Mean
C. Standard devistion

Three formulas are offered as suggestions:

$$\sigma=\sqrt{\frac{\Sigma f(x^2)}{N}}\ ;\ \sigma=\sqrt{\frac{\Sigma f(d^2)}{N}-\left(\frac{\Sigma fd}{N}\right)^2}\ ;\ \sigma=C\sqrt{\frac{\Sigma f(d')^2}{N}-\left(\frac{\Sigma fd'}{N}\right)^2}$$

You may use the calculator or any books or notes handy (in this office) to assist you.

The scores are fictitious:

22
45
63
88
55
12
45
25
36
96
98
52
52
12
45
23
23
56
48
19

15
23
15
42
78
45
26
13
12
53
50
60
88
91
50
60
55
48
75
42

13
56
48
59
58
45
75
36
26
15
24
25
85
76
13
23
56
45
78
15

24
35
37
31
25
16
92
53
56
65
23
34
68
97
65
34
39
36
58
53

26
29
36
57
58
35
64
67
12
95
63
68
45
78
76
16
23
58
59
49

75
76
16
72
42
48
59
67
37
38
24
35
19
25
85
74
85
46
63
52

41
74
14
25
36
78
56
23

The preparation of a test or scale is like writing an important speech or research report for publication. The first draft is completed after many hours of research, reading, writing and re-writing. Finally, an experimental draft is prepared, after a pilot study, and the instrument is administered to an experimental group frequently numbering in the thousands. The results are then analysed to determine whether or not changes should be made. Assume all of the above procedures have been completed and you are at this stage of development in the instrument construction. Use your project.

What would you do now?

Be very specific and illustrate wherever possible such as how you would word the changes, etc.

A. General format
B. Instruction for administering
C. Answer form
D. Item content or wording
E. Scoring
F. Sample respondents
G. Analysis of results, standardization
H. General treatment of the topic or subject. (Do not suggest a new topic or subject--assume your assignment was to construct an instrument just as you chose.)

Following are examples of test items prepared by students concerning the subject matter of Chapter 23 in the text. All items are in the form of definitions. Each was prepared to include one correct response. Choose the best response (if one does exist) and prepare a third response that is not a proper response but is plausible. If no correct response is included in the example, prepare the third response as the correct response. Try to avoid changing the wording of the opening statement but, if in your opinion, a proper response can not be made to the statement, change the statement and re-write all three responses. The textbook and the dictionary may be used as you see fit. Be sure to clearly indicate the correct response. Select any fifteen.

1. Variability could best be classified as:
    a. either reliable or unpredictable
    b. inconsistent

2. In unreliable instruments, the addition of items of equal kind and quality will be more likely to produce:
    a. greater systematic variance
    b. greater reliability

3. The square of the coefficient of correlation is:
    a. the coefficient of determination
    b. the reliability coefficient

4. All valid scores most likely include:
    a. an error component
    b. the true score

5. The first and most elementary step in any measurement procedure is to:
   a. define the objects of the universe of discourse
   b. define the properties of the objects of the universe of discourse

6. In the absence of sufficient criteria, when we say two objects are the "same" we mean:
   a. they are identical
   b. they are sufficiently the same to be classed in the same set

7. An example of a transitivity postulate is:
   a. $A > B < C$; $A \geq C$
   b. $A > B > C$; $A > C$

8. The range in measurement is:
   a. the objects being measured
   b. numerals assigned to the objects being measured

9. A construct is:
   a. an operational definition
   b. an invented name for a property

10. The measurement ideal is:
    a. the ratio scale
    b. the equal-interval scale

11. If (a=b) and (b=c), then (a=c); this postulate expresses:
    a. transitivity
    b. equality alone

12. A construct is:
    a. usually arbitrary
    b. the same as "the reason for"

13. A function is:
    a. a relation in which with a given first element, there are never two different second elements
    b. a relation in which two ordered pairs have the same first element and different second elements

14. An ordinal number:
    a. tells the position of an object in a particular linear sequence or succession
    b. indicates absolute quantities which are positioned according to their total value.

15. Equal-interval scales are characterized by:
    a. equal distances on the scale representing equal distances in the property being measured
    b. measurement taking place at equal intervals of time

16. In which of the following does the word "figure(s)" refer to a numeral?
    a. "This measurement has three significant figures."
    b. "The figure in the ten's place is wrong," said the teacher.

17. Nominal measurement:
    a. assigning symbols which have qualitative but not quantitative meaning to sets of objects
    b. is estimating the actual measure of something

18. An action taken to infer some unobservable characteristic is:
    a. an indicant
    b. a construct

19. A postulate is:
    a. an unproven law accepted on faith, unrelated to theorems
    b. an assumption necessary to forward movement of a project

20. Measurement is the:
    a. assignment of numbers to objects or events according to rules
    b. assignment of numerals to objects or events according to rules

21. Ordinal numbers:
    a. tell "how many"
    b. tell "which one"

22. A division of the universal set into subsets that are disjoint and exhaustive is:
    a. a partitioning of U
    b. a means of limiting the variables of U

23. A cardinal property of a number is that property that a math set has in common with all sets that can be put into a:
    a. 1 to 1
    b. 1 to 2 correspondence

24. Measurement in the broadest sense is:
    a. the assignment of numerals to objects or events according to rules
    b. the accessment of values given to any object or event

25. The classes El and Ell correspond to:
    a. the domain
    b. the range

26. Superior intelligence, bright and high achiever

In the above example the word indicant applied to:
    a. "superior intelligence, bright and high achiever."
    b. his class standing and scores test situations

27. Superior intelligence, bright and high achiever

Using the above example the word construct applied to:
    a. "superior intelligence, bright and high achiever"
    b. John's grades and the fact that he is in the upper 3% of his class

28. Partitioning a set consists of:
    a. breaking it down into subsets that are mutually exclusive and exhaustive
    b. defining the objects of the universe of discourse

29. Dichotomous:
    a. having only 2 values
    b. having various subsets

The subject of tests and measurements is characterized by a differentiation between reliability and validity. The end product of the subject is a matter of inference. Discuss the relationship that exists among these three concepts.

Write an opening statement and a series of at least three funnel questions that could be used in an interview schedule leading up to the question: "How did you receive improper help on the examination you were given in the chemistry final exam?"

Select another member of the class for observation and during the course of the evening, record at least ten observations of that person from which one could infer a behavioral construct.

Evaluation Criteria
Final, **Education 4512**

Instructions: Mark each answer in the space provided at the left margin of the paper or on the line provided. Questions requiring written comment may be continued on the reverse side of the question sheet. Be sure your name appears on each page. Unless otherwise stated, pupils are considered to be the target population.

Questions 1-15: Three research studies (A, B, and C) are briefly described. Following each study is a series of five test items asking you to indicate the type of study followed by a series of statements for which you are to indicate whether the conclusions stated are permissible or not permissible. Use "P" for each permissible statement and "NP" for each not permissible conclusion. Assume that differences reported under "Findings" for each study are statistically significant. Assume also that the sample of subjects for each study is representative of the population about which the statements are made.

* * *

Study A.
Question: Is the number of counseling sessions with high school counselors related to student matriculation in college?

Procedures: Researcher counts number of counseling sessions recorded (the records are accurate) on each subject's completed high school Counseling Record Card and compiles an accurate record of the number of subjects who matriculate in college.

Findings: The greater the number of counseling sessions, the greater the number of students matriculating in college.

* * *

_____ 1. Indicate whether this is a status study, an associational study, or an experimental study.

_____ 2. The number of students matriculating in college will increase if each student is given an increased number of counseling sessions. P or NP.

_____ 3. More counseling sessions result in greater student matriculation in college. P or NP

_____ 4. There is a positive relationship between number of counseling sessions and student matriculation in college. P or NP

_____ 5. Some students matriculate in college because of having an increased number of counseling sessions. P or NP.

* * *

Study B.
Question: Is the type of advice about college attendance that a counselor gives to the student related to student matriculation in college?

Procedures: Researcher randomly assigns one-third of all high school seniors to Group X, one-third to Group Y, and one-third to Group Z. Counselors tell each Group X student about college and advise each student to attend. Counselors tell each Group Y student about college but do not give advice about whether or not the student should attend. Group Z receives no counseling or advice about college. Records are maintained to indicate which students from each group matriculate.

Findings: 450 Group X students matriculate
300 Group Y students matriculate
300 Group Z students matriculate

* * *

_____ 6. What type of study is this? (Status, associational, or experimental)

_____ 7. Counseling that includes "attend college" advice to the student results in an increase in the number of students matriculating in college. P or NP

_____ 8. More students advised by counselors to attend college matriculate than do students not given counselor adivce about whether or not to attend college. P or NP

_____ 9. Some students matriculate in college because they receive counselor adivce to attend college. P or NP

_____10. The number of students matriculating in college will be greater if students are advised to attend college than if they are not given much advice. P or NP

* * *

Study C.
Question: Is the number of students matriculating in college related to socio-economic status (SES)?

Procedures: In the same high school, 100 students are randomly selected from high SES, 100 from medium SES, and 100 from low SES, A record is maintained to indicate which students from each group matriculate in college.

Findings:
82 High SES students matriculate
54 Medium SES students matriculate
27 Low SES students matriculate

* * *

_____11. What type of study is this? (Status, associational, or experimental)

_____12. Changing the social class of a group of 100 students from low to medium will increase the number of students from that group who matriculate in college. P or NP.

_____13. Socio-economic status is positively related to matriculation in college. P or NP.

_____14. Some students matriculate in college because they have high socio-economic status. P or NP.

_____15. The higher the socio-economic status of a given group of students, the greater the probability that students from the group will matriculate in college. P or NP.

In the following indicate your choice of reply by indicating the number in the space provided. Select the best reply for each item. Use "0" for "none of these".

_____16. Which variable in the three studies above is an independent variable?
1. Number of counseling sessions
2. Student matriculation in college
3. Socio-economic status
4. Type of advice about college attendance

_____17. Which variable in the three studies above is a dependent variable?
1. Number of counseling sessions
2. Student matriculation
3. Socio-economic status
4. Type of advice about college attendance

_____18. All of us are concerned these days about children's problem-solving abilities. Great claims have been made about the potency of the newly developed Mental Flexibility Training Program in improving problem-solving achievement as measured by standardized problem-solving tests. You therefore design an experimental study to compare the Mental Flexibility Training to the Ordered Thinking Program (which your district currently uses) in terms of their effects on problem-solving achievement. What is the dependent variable in your study?
1. Type of training program
2. Mental Flexibility Training
3. Problem-solving achievement

_____19. What is the independent variable in the study described in item 18?
1. Type of training program
2. Mental Flexibility Training
3. Problem-solving achievement

_____20. Which one of the following is not usually considered a threat to the internal validity of an experiment?
1. History
2. Reactive arrangements
3. Selection
4. Mortality

_____21. Suppose you have two intact classes available, both taught by the same teacher, and wish to test the efficacy of a particular visual aid in instruction. For each unit through the semester the aid is assigned randomly to one class or the other. At the end of a semester, the units taught by visual aid are significantly superior to those taught without. You can generalize the results to presence of the aid in general under which of the following assumptions?
1. No interaction exists between treatments and any of the specific variables such as teacher, or class, etc.
2. This teacher is an average teacher and the class is an average class.
3. No interaction exists between maturation and history.
4. The classes, although intact, are essentially equivalent.

_____22. When the results of an experiment are true only under certain very special conditions and do not generalize far beyond the particular experimental conditions, the experiment is said to have

1. low internal validity
2. high internal validity
3. low external validity
4. high external validity

_____23. Which one of the following is usually considered a threat to the internal validity of an experiment?

1. Interaction of testing and treatments
2. Multiple treatment interference
3. Regression
4. Reactive arrangements

_____24. One is interested in ascertaining if distribution of an exercise pamphlet in a secretarial pool causes the weight of each woman to decrease. Each woman was weighed before and after the distribution of the pamphlet. On an average, each woman's weight had decreased. Therefore, the conclusion was that the exercise pamphlet induced the women to lose weight. This example illustrates the experimenter's neglect of which threat?

1. History
2. Regression
3. Testing
4. Testing and X-interaction

_____25. In some speech studies of early childhood, the procedure was merely to listen to the spontaneous vocalizations of an infant and write down what was heard. The selective factor of auditory perception--listeners "hear" most readily those sounds that correspond to the phonemes of their own language--was not considered. What threat did the experimenter not give attention to?

1. Maturation
2. Instrumentation
3. Multiple-X interference
4. History

_____26. It is possible to obtain internal validity without having

1. reliability
2. generalizability
3. replicability
4. variability

DIRECTIONS: Continuing to mark the correct answer on the test sheet, use the following designs to answer the questions below. X is some experimental treatment. R is the random assignment to groups. 0 is some observation of behavior. M is equivalent materials. A broken line indicates the use of intact groups without random assignment.

1. $0_1 0_2 0_3 0_4$ X $0_5 0_6 0_7 0_8$

2.
```
0 0 0 0 X 0 0 0 0
- - - - - - - - -
0 0 0 0 X 0 0 0 0
- - - - - - - - -
0 0 0 0 0 0 X 0 0
```

3.
```
        R
0 0 0 0 X 0 0 0 0
- - - - - - - - -
0 0 0 0   0 0 0 0
```

4. $X_1 0$ $X_0 0$ $X_1 0$ $X_0 0$ $X_1 0$ etc.

5. $M_a X_1 0$ $M_b X_2 0$ $M_c X_1 0$ $M_d X_2 0$ $M_3 X_1 0$ etc.

6.
```
0 X 0

0 - 0
```

7.
```
R 0 (X)
R   X  0
```

8.
```
R 0 (X) (0)
R    X   0
- - - - - -
R 0
R        0
```

9.
```
Present class X old 0
               - - - - - -
Next year class X new 0
```

_____27. The particular advantage of design (3) over design (1) is that some attempt is made to control
1. history
2. selection
3. mortality
4. multiple-X interference

_____28. The advantage of design (8) over design (6) is the comparative absence in design (8) of the danger of
1. instrumentation
2. interaction of selection and X
3. multiple-X interference
4. interaction of testing and X

_____29. Where one cannot control the treatment but can control the observation random assignment to the measurements, then one may wish to use which design of the panel?
1. Design 1
2. Design 2
3. Design 7
4. Design 8

_____30. A teacher was interested in testing the value of having students use flash cards while learning new vocabulary words. She has only one class of thirty students to test. What design would be best?

1. Design 1
2. Design 5
3. Design 7
4. Design 10

_____31. Your school district has been using a team teaching approach in ninth grade English for three years in each of eight high schools. The superintendent wishes to know whether team teaching results in any greater learning than a traditional approach. He makes the resources of the Research Center available to you and obtains permission for you to use ninth grade students in a nearby high school. These students are being taught ninth grade English by a traditional approach. The experiment is to run during the school year 1967-68. Choose the best design.

1. Design 4
2. Design 6
3. Design 8
4. Design 9

_____32. A junior high school has had a probelm in pupil behavior. It is associated with pupil attitudes. The school institutes a series of discussion groups with representatives of each class participating on several occasions. Each child will eventually participate five times. You will frequently measure the ongoing effectiveness of the progeam in changing pupil attitudes. Which design is best?

1. Design 1
2. Design 5
3. Design 6
4. Design 8

_____33. A retention study is to be made with students who took BSCS Biology. About 750 sophomores (approximately 86 from each of 9 schools) took BSCS and are still in attendance as juniors. Data are available on these students and an equal number of sophomores who took traditional biology. What is the best design?

1. Design 2
2. Design 6
3. Design 8
4. Design 9

_____34. A sixth grade teacher wants to see what effect, if any, fifteen minutes of vigorous exercise just prior to a reading lesson has on reading performance. The other three sixth grade teachers are willing to use the fifteen minute exercise. Which design is best?

1. Design 1
2. Design 2
3. Design 5
4. Design 7

35-40. Mark 1 on your test sheet for any of the following instructional objectives which are properly stated. Mark 2 for any which are improperly stated.

_____35. The student will develop interest in leisure sports

_____36. The student will accurately learn the best known works of Voltaire.

_____37. The teacher will help the class to solve algebra problems correctly.

_____38. The student will appreciate the key importance of algebraic approaches.

_____39. The student will include ten supporting facts in a written persuasive paragraph.

_____40. The student will become familiar with how to write an essay using no reference but personal experience.

* * *

41-43. Mark 1 on your test sheet for each of the following that indicates a change in educational means. Mark 2 for each item that indicates a change in ends.

_____41. Use new self-study pamphlets to teach current math topics.

_____42. Drop biology course objectives dealing with sex education.

_____43. Use increased quantities of teacher praise in classroom discussions.

* * *

44-46. Mark 1 for each dependent variable, 2 for each independent variable.

_____44. The quality of students' essay examinations.

_____45. Frequency of teacher's efforts to show pupils the value of the subject matter.

_____46. Amount of time teacher threatens pupils with possible failure in the class.

* * *

47-49. Mark 1 for each manipulable variable, 2 for each non-manipulable variable.

_____47. Number of pupils in class.

_____48. Socio-economic class of pupil.

_____49. Parental level of education.

* * *

50-52. Mark 1 for each of the following treatments that is operationally defined. Mark 2 for each treatment that is not.

_____50. Use of the Encyclopaedia Britannica Films, Inc., Harvey White Physics Series in all "honors" physics class. Each daily film to be followed by fifteen minutes of class discussion led by the teacher.

_____51. Use of a four item multiple-choice review at the beginning of each class period.

_____52. An individually differentiated instructional scheme in which each learner is given instructional tasks consonant with his intellectual ability and experiential bakcground.

* * *

53-57. Mark 1 for each item below that could be employed as a criterion to evaluate educational programs. Mark 2 for each item that could not be used as an educational criterion measure.

_____53. Tape recordings of how individual pupils respond to a surprise question by the school's vice principal: "What do you want to be when you grow up?"

_____54. A listing of the non-school clubs voluntarily joined by students during their senior year in high school.

_____55. School attendance records.

_____56. The age of learners.

_____57. Locally constructed tests of pupils' progress in spelling.

* * *

58-63. Using the eight category scheme presented below, classify each of the following numbered items by writing the appropriate number in the space provided on the test sheet.

1. Learner-behavior - natural stimuli - locally devised
2. Learner-behavior - natural stimuli - commerically devised
3. Learner-behavior - manipulated stimuli - locally devised
4. Learner-behavior - manipulated stimuli - commerically devised
5. Behavior product - natural stimuli - locally devised
6. Behavior product - natural stimuli - commerically devised
7. Behavior product - manipulated stimuli - locally devised
8. Behavior product - manipulated stimuli - commerically devised

_____58. An examination of the annual tonnage of school litter found in the halls as an index of school citizenship.

_____59. Inspection of student reading preferences by the amount of monthly wear on magazines available in classrooms during free reading periods.

_____60. Use of Flanders' observation schedule for recording classroom behavior.

_____61. School attendance records.

_____62. California Achievement Test Scores

_____63. Pupil responses to teacher-made questionnaire regarding their normal home relationships with parents and siblings.

* * *

64-65. Write in operational terms two objectives that could be derived from the following educational objective: "To understand how climate and georgraphic conditions have exerted an important influence on the social and economic development of man." Answer in the space provided on the back of this sheet.

66-68. List and describe (in one sentence each) three procedures which can be employed to increase the probability that the requirements of the experimental treatments are met in classroom research and/or to determine whether the requirements were met. Answer on the back of this sheet.

APPENDIX B

Examples of Team Projects

THE EFFECT OF THE PARTICULAR WEEK DAY OF TEST ADMINISTRATION ON NINTH GRADE STUDENT ACHIEVEMENT AS MEASURED BY DAILY TESTING

A RESEARCH STUDY CONDUCTED
FOR
EDUCATIONAL TESTS, MEASUREMENTS,
AND PRACTICUM APPLICATIONS II
EDUCATION 4512

BY

LEA W. JOYNER
C. H. HOLLAND

with assistance from Austin Welsh

MAY 1968

ACKNOWLEDGEMENTS AND APPRECIATIONS

The researchers on this experimental study wish to take this opportunity to extend their heartfelt appreciation and a special word of gratitude to Mr. Gordon Gilbert, principal at Bellevue Junior High School, for his ready approval and willing cooperation on this study and throughout the student teaching practicum of this writer. Without his cooperation, implementation of this project would not have been possible.

The research team also wishes to express their appreciation and gratitude to Mrs. Jacobs, the librarian, who also teaches in the evening division at Memphis State University and Mrs. Hickock, secretary to the principal, who was helpful in so many ways with the "little things."

We would like to remember also the professors and advisors at Memphis State University whose time was so generously given in aid and assistance to these novice investigators.

Last but not least — a very special appreciation is extended to each of the one hundred thirty nine wonderful children who participated in this pilot study — albeit unknowingly.

They were the greatest!!

PREFACE

There are, at the fingertips of the classroom teacher, innumerable opportunities for research which may be the source of additional knowledge and insight into many of the problems of understanding and advancing the teaching-learning process. While everyone must operate within the limits of his or her teaching assignment, level of training in research techniques and necessary content background, the teacher can and should become skilled in conducting quality research in many areas. Among the principal points of justification for such endeavors, professional growth and knowledge, problem solving and systematic theory building are especially worthy of consideration. But whatever purpose the research is to serve, a statement of the underlying philosophy and rational is indispensable. A systematic, thoughtful and vigorous approach to the investigation is necessary if we are to expect a clear design and a valid strategy of attack. Only when tight and organized frames of reference are consciously employed will the field of education move toward quality research.[1]

---

[1] Byron G. Massialas and Frederick R. Smith, "Quality Research — A Goal For Every Teacher," Applied Research in Education. Totowa, New Jersey: Littlefield, Adams & Co. 1965. Pp. 16-17.

INTRODUCTION

There were thirteen senior students selected to participate in the Undergraduate Research Training Program at Memphis State University in the Fall-Spring Semester of 1967-68. The program, funded by the United States Office of Education under Title IV of the Elementary and Secondary Education Act, held as its primary objective the development of research awareness and statistical ability in the students.

The first semester was devoted to:

1. Developing basic skills in statistical procedures.
2. Development and administration of test instruments.
3. Familiarization with proper methods for searching the literature.
4. Development of an awareness of the threats to validity present in most research studies.

These courses were under the pedagogical direction of Dr. Donald R. Thomsen, Dr. Jimmie C. Fortune and Dr. Jack Miller.

At the beginning of the second semester the students enrolled in Education 4512 (Educational Test and Measurements Practicum) and soon thereafter embarked upon the practicum phase of the program under the supervision of Dr. Thomsen. The class members were instructed to conceive, design, develop, implement and analyze an experimental research study to fulfill a major segment of the course requirements. The class was divided into teams of two to four members and required to select their own areas of research independent of faculty advice or suggestion. L. W. Joyner and C. H. Holland formed a two member team and proceeded to search for a practical and feasible area in which they would be able to conduct a study.

Joyner, concurrently engaged in his student teaching practicum at Bellevue Junior High School, was serving under Mr. Austin Welsh in ninth grade General Science and Miss Mary Jordan in seventh grade Tennessee History.

During the course of a discussion related to an upcoming test the questionable validity of a Monday testing situation was mentioned by Mr. Welsh. After considerable discussion a preliminary searching of the literature was conducted; this effort failed to reveal any available research to substantiate claims toward any viewpoint relating to this area of activity. It was thus decided that a potential for limited but quality research on this topic did indeed exist. Based on these preliminary conclusions, the investigators decided to write a proposal for a related research project to be conducted in this area and set forth the limits that would be imposed on such an investigation. A study based on the determination of any relationships that might exist between student achievement and the week-day of test administration was decided upon.

The proposal, written by L. Joyner and C. Holland, was presented to Mr. Austin Welsh, science teacher, and Mr. Gordon Gilbert, principal. Mr. Welsh not only approved the study but also wrote an endorsement to Mr. Gilbert requesting his very necessary approval which followed within twenty-four hours.

During the course of the investigation Mr. Welsh proved to be an indispensable member of the team not only acting in an advisory capacity, but also very actively contributing by the daily administration of the test and the carrying out of other tasks. Without his full cooperation and assistance this study would not have been possible.

CONSENT FOR THE STUDY

During the time this writer was assigned to Bellevue Junior High School, for the purpose of performing his student teaching practicum, there was also a pressing need to secure permission to conduct an experimental study in some area related to educational research. The latter was a necessary requirement for Education 4512 under Dr. Donald Thomsen, Associate Professor, Memphis State University, Education Department and Assistant Director, Educational Information Systems, Central Midwestern Regional Educational Laboratory, Inc., 3808 Norriswood, Memphis, Tennessee.

When the present area of study was first suggested by Mr. Austin Welsh, it was immediately given serious consideration by the concerned individuals and quickly adopted as the study to be conducted by these workers.

A proposal for a research study was written and submitted to the following authorities for their necessary approval:

NAME	POSITION	ACTION TAKEN
Dr. Donald R. Thomsen	Associate Professor Memphis State University, College of Education and Assistant Director Educational Information Systems, Central Midwestern Regional Educational Laboratory, Inc. 3808 Norriswood Memphis, Tennessee 38111	Approval
Dr. Jimmie C. Fortune	Assistant Professor, Memphis State University College of Education , Bureau of Educational Research and Services Memphis, Tennessee	Approval
Dr. Mosely Powell	Assistant Professor College of Education and University Supervisor for Student Teaching Practicum Memphis State University Memphis, Tennessee	Approval
Mr. Gordon Gilbert	Principal Bellevue Junior High School Memphis, Tennessee	Approval
Mr. Austin Welsh	Instructor in General Science Ninth Grade Bellevue Junior High School Memphis, Tennessee	Approval

The investigators named herein proceeded to conduct the study according to the general guide lines, requirements and limitations as set forth in the above named proposal and more fully described in this paper.

CLIMATE—ENVIRONMENT AND LEARNING ATMOSPHERE

Bellevue Junior High School is a forty-one year old public educational facility in the Memphis City School System comprising grades seven, eight and nine. The original red brick structure has undergone limited modifications and additions since its erection in 1927. For a building of this age, the state of repair is good. The custodial and maintenance services furnished by the city board of education are adequate.

Teaching facilities appear adequate to meet minimal student needs, but there is a significant lack of equipment in certain areas. There are several film projectors but very few television sets are available. This is true of most other types of audio-visual equipment also. The one facility short coming most closely related to this project was the absence of adequate laboratory facilities for the one hundred thirty nine students observed during the course of the study. The project was programmed around this deficit, however, and no major problems presented themselves in this area.

THE STUDENT BODY

The student body at Bellevue Junior High School is composed of six hundred eighty students, racially mixed, with approximately twenty percent Negroes throughout the classes.[2] The transfer ratio is significantly higher than for most other junior high schools in the city.

SOCIO-ECONOMIC BACKGROUND

The socio-economic background of the student population has declined steadily in recent years and at present a relatively low socio-economic status is in evidence for a large number of the children. Bellevue was rated as a deprived area school last year and the principal and faculty feel that it should be so rated at this time. Additionally, the sexual distribution of the student population has recently shifted from a positive male balance to a negative male balance, and the average age levels of these minority male and majority female groups have increased.

FACULTY AND OTHER PERSONNEL

The staff is currently composed of twenty-seven faculty members, two guidance counselors, administrative personnel, one librarian plus cafeteria and custodial personnel. Ages vary and appeared to this observer to be about evenly distributed along a scale from early twenties to almost retirement age. The faculty is predominately white and female. There are two Negro teachers, one male and one female. Ages of these two instructors were estimated by this observer to be middle or late twenties. There seemed to exist a healthy and cooperative relationship between the latter and the rest of the faculty. Tensions were not obvious and no outward feelings of a racial nature were observed.

---

[2] This is an "official" estimate and according to Mr. Gilbert varies because of the high transfer ratio at this particular school. This percentage figure was more or less approximate throughout the five ninth grade General Science classes involved in this study.

OTHER BACKGROUND INFORMATION

These workers feel that, generally,the rapport established with the children in the classes under observation was good to excellent. It is also felt that a major contribution to the excellent relationship with the students may be directly related to the student teaching practicum that preceeded the actual study.

OBSERVATIONS REGARDING CLASS ATTITUDES

Most of the students in the ninth grade General Science classes observed during the course of this investigation did not demonstrate any exceptional enthusiasm for the science material studied. Indications were that they took the course merely to fulfill the credit requirement at this level. There were a few exceptions in each class and these particular individuals usually performed at levels well above class norms.

Attitudes were generally good and, as previously stated, the rapport between students, teachers and researchers was excellent. This positive relationship was especially noticeable in Mr. Welsh's classes. According to unofficial reports and "educated guesses" among the faculty, the general performance of the majority of the student groups at Bellevue was usually below expected norms for similar groups city-wide. No substantiating evidence was presented regarding this rather subjective observation by faculty and it is mentioned here merely as an indicator of teacher attitudes.

HYPOTHESIS

As stated previously, this study proposed to produce some evidence to indicate an optimum day of the week for classroom evaluation and testing. The working hypothesis for the experimental effort was therefore: THERE IS A SIGNIFICANT RELATIONSHIP BETWEEN STUDENT ACHIEVEMENT AND THE WEEK DAY OF TEST ADMINISTRATION. The investigators further expected to obtain a definite focus of data toward one particular day of the week.

STATEMENT OF THE PROBLEM

It is the understanding of these investigators that there are some students of human behavior, among whom, the concept of daily peaks in efficiency and performance in the individual is generally accepted. The importance of learning and testing activities that would be associated with any possible extension of this concept from an individual-hour of the day basis to a group-day of the week basis is obvious.

If such an extension were found to be valid and permissable then certain new approaches to learning situations, as well as most classroom measurement and evaluation, may be indicated and should certainly be considered from a pedagogical position. Since one of the major concerns of educators is achievement by the student, it follows that awareness by the instructor of an existing optimum week day for evaluation and testing would be a necessary and vital step toward improved student performance in most areas of learning. A related consideration would be the converse regarding a particular week day that might indicate minimal results or achievement.

The selection of a definite day of the week for evaluation and testing in any particular discipline has, in the past, been based primarily on custom, tradition and teacher convenience. In many cases it has become an accepted practice to group testing situations into certain periods of time; such as, six weeks, semester and final examinations to coincide with administrative periods in the academic year. Within this structure, however, teachers are given certain latitudes in which to operate. Some teachers, for example, may not give a six weeks test but may compute grades from previous measurements. Others may give only six weeks examinations and so on.

The study described in this paper was concerned with gathering evidence in order to facilitate the establishment of norms or standards for making decisions related to test administration based primarily on a day of the week concept.

One of the primary problems with which these investigators concerned themselves was related to the almost universal aversion to administration of tests on Mondays.

As indicated elsewhere in this report, a search of the literature revealed no evidence of any research or controlled investigation to substantiate such a policy. It is therefore concluded that the criteria for such practices are:

1. Tradition and custom
2. Uncontrolled classroom observations
3. Teacher convenience and arbitrary selection.
4. Hearsay and "experience."

These situations are subject to all the threats to validity and reliability offered by such uncontrolled non-scientific methods of observation and measurement.

This study proposes to determine whether or not there is an optimum day of the week for classroom evaluation and testing, and if so to indicate that measurements obtained on this day will reflect improved student achievement.

THE RESEARCH DESIGN

After extended consultations with Dr. Thomsen and Dr. Fortune regarding the model for the research design it was decided to use a daily testing situation incorporating short item quizzes into a time series study in which the day actually became the sample observed. The data collected was then subjected to a sequential analysis of variance or an ANOVA on the same cell recorded over the remaining days of the week.

The administrative problem of separating the lecture or presentation of content material from the test construction and recording was solved with the assistance of Mr. Welsh who presented Joyner a set of objectives which served as a guide to the selection of the test items. Mr. Welsh then administered the tests while Joyner developed the test items, scored and recorded the grades.

With minor modifications the design used in this study was constructed generally along the lines of what Campbell and Stanley term COUNTERBALANCED designs. Counterbalanced designs are quasi-experimental designs and include all efforts in which experimental control is achieved, or precision enhanced, by entering all respondents, or settings, into all treatments.

> ... such designs have been called "rotational experiments" by McCall (1923), "counterbalanced designs" (e.g., Underwood, 1949) "cross-over designs (e.g., Cochran and Cox, 1957; Cox, 1950), and switch-over designs (Kempthorne, 1952)...[3]

The design is portrayed in the following diagram with post-test only because it is especially preferred where pretest are unavailable or inappropriate as in the particular study described herein.

(WEEK 1 OF 7)

	MON	TUE	WED	THUR	FRI
Group A (Class Period #1)	$\underline{X}\ \underline{O}^R$	$\underline{X}\ \underline{O}^R$	$\underline{X}\ \underline{O}^R$	$\underline{X}\ \underline{O}^R$	$\underline{X}\ \underline{O}^R$
Group B (Class Period #2)	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$
Group C (Class Period #3)	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$
Group D (Class Period #5)	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$
Group E (Class Period #6)	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$	$\underline{X}\ \underline{O}$

(AND SO ON FOR 7 OF 7 WEEKS)

R = Random assignment of pupils to comparison group.
X = Exposure of students to an experimental condition.
O = Some form of observation or measurement.
Broken Line = the broken lines between the lines signify the individuals in the groups were not chosen at random; thus, the groups are not equivalent.

---

[3] Donald T. Campbell and Julian C. Stanley, Experimental And Quasi-Experimental Designs For Research, Chicago: Rand McNally & Company, p.50. 1966.

The coding here has specific meaning. Whenever X's and O's appear on a given line the application is made to the same specific person. X's and O's that are in a vertical relationship refer to a simultaneous same day application but to another person or persons. Randomization within groups permits the assumption the groups are equivalent. It was not possible, however, to achieve randomization in this study. Therefore, the groups are not equivalent.

In this design the differences in sums could not be interpreted simply as artifacts of the initial group differences or of practice effects, history, etc. Similarly comparable are the sums of the rows for intrinsic group differences, and the sums of the columns of the first presentation for the differences in occasions.[4]

For the purposes of this study the counterbalanced design described by Campbell and Stanley, was modified so that the occurrence of each successive Monday, Tuesday, Wednesday, Thursday and Friday became the sample for the experiment. The data recorded were the individual test scores and the class-period mean of these scores for each testing week day became the prime observations for the experiment.

The experimental treatment consisted of the daily administration of the five item quiz on each occuring week day.

For purposes of this study the class-period groups were non-equivalent. They were not ability grouped by the school prior to the study described herein. As in all quasi-experiments, strength was gained through consistency of internal replications of the experiment as time allowed.

The investigators were prone toward a design of the type described here because of the lack of scheduling control over the relatively few naturally aggregated groups of the classes. Unable to subdivide these natural groups into randomly equivalent subgroups for either presentation of experimental treatment or testing, and after consultation with Doctors Thomsen and Fortune, the investigators became fearful of threats related to interaction of selection and occasions. The possibility of such an occurence was deemed less likely in the modification of the counterbalanced design described herein, because all comparisons are demonstrated in each group, over each day, over each week and hence several matched interactions would be required to initiate the experimental effect. In other words, any such coincidental effects or threats would have to occur on separate occasions, over repeated times, in each of the groups in turn.

---

[4] Donald T. Campbell and Julian C. Stanley, Experimental And Quasi-Experimental Designs For Research, Chicago: Rand McNally & Company, p. 50. 1966.

DEFINITIONS AND LIMITATIONS

For purposes of this study student achievement criteria was mastery of a limited number of concepts and factual knowledge. Standard grading techniques were used.

The test day was that particular day of the school week in which the testing instrument was administered to the sample groups. Briefly, five-item daily quizzes were used.

The test instrument was considered as that instrument designed to measure student accomplishment and performance in the particular unit of study under consideration for each testing period. The applied instrument was composed of written questions, short answers, true--false and completion. It evaluated acquisition of factual knowledge and understanding of concepts related to scientific subject matter at the ninth grade level.

No effort was made to standardize or establish norms beyond those generally accepted for the material and grade level.

The test made no specific effort to measure previous or general knowledge.

The experimental treatment was defined as the week day of test administration and was presented on a daily basis.

For the particular purposes and needs of this study the individual student score was an item while the arithmetic mean for the total group was the observation recorded as data.

Under the conditions and the design of this study the sample was considered to be the actual week day of test administration and occurred five times per class per week, with exceptions noted.

TEST DESIGN AND CONSTRUCTION

1. The design and construction of the test was based on an outline of objectives as presented to Joyner by Mr. Welsh.

2. The individual items were placed on index cards—one item per card. Answers and page references in text were placed on the card in parenthesis.

3. The source material for all evaluations was the text book for this course as used in the Memphis City School System: Willard J. Jacobson, Robert N. King, Louise E. Killie, and Cecilia J. Lauby, Challenges in Science (second edition), New York: American Book Company, 1964.

4. One hundred eighteen pages of material from units seven and eight in the text formed the content basis for the evaluation instruments used in the experimental treatment as applied to this investigation. These seven chapters, twenty-three through twenty-nine, were primarily concerned with human behavior and science in the home and community.

5. The test design attempted was more achievement centered than aptitude centered, although the distinction between these two constructs is by no means definite. In addition to being a measuring device it was also desired that the tests serve a motivational purpose that would reflect positively for the instructional phase of the study. The scoring was thus used as a method for immediate reinforcement by allowing the children to grade their own papers.[5] Problems arose in this area that are more fully discussed elsewhere in this paper.

6. Most multiple-choice and short answer items were constructed according to recommendations by D. A. Wood.[6]

7. Other elements considered important during test construction were:[7]
    - A. Clarity of expression.
    - B. Brevity of expression.
    - C. Avoiding irrelevant sources of difficulty.
    - D. Technical terms as specific determiners.
    - E. Avoiding irrelevant inaccuracies.

8. The students recorded all answers to questions on a small quarter section of paper approximately four inches wide by five and one half inches long. Only the answers were written by the students.

---

[5] Due to the heavy work load of the research team members, this technique also served as a time saving device.

[6] Dorothy Adkins Wood, Test Construction, Columbus, Ohio: Charles E. Merrill Books, Incorporated, 1961. Pp. 112-25.

[7] Ibid., pp. 44-50

SUPPLIES AND EQUIPMENT

The only extra supplies and equipment necessary for this study were small three by five index cards supplied by the investigators. The tests were taken on the reverse side of old unused mimeographed class handouts gathered together for this purpose by Mr. Welsh. These sheets were cut into four equal sections and distributed to the classes. Therefore, supplies and equipment necessary for this study were furnished by the above named investigators at no expense to Bellevue Junior High School or the city school system.

DATA GATHERING ACTIVITIES

At approximately 8:50 on the morning of March 25, 1968 the actual classroom investigation for this study was initiated among one hundred thirty nine, ninth grade, General Science students taught by Mr. Austin Welsh at Bellevue Junior High School, Memphis, Tennessee.

The actual experimental treatment was terminated at approximately 2:30 P.M. on May 10, 1968 after having been conducted daily, Monday through Friday for a total of seven weeks in all five science classes. During this period of time the investigators were able to accumulate data for twenty-nine days of actual test administration.

The sum of all data gathering activities over the seven weeks duration of this study resulted in a total of 3,473 individual experimental treatments on the following basis:

PERIOD	TEST ADMINISTRATIONS (INDIVIDUAL)
First Period	740 Test Administrations
Second Period	786 Test Administrations
Third Period	763 Test Administrations
Fifth Period	601 Test Administrations
Sixth Period	583 Test Administrations

Another point of interest to this investigator was the discovery, upon initation of the statistical treatment, that during the course of this investigation, based on a 1:1:1 relationship a student confronted an item on a day for a Grand total of over seventeen thousand student-item-day encounters* according to:

FIRST PERIOD:	3700	Student-item-day encounters
SECOND PERIOD:	3930	Student-item-day encounters
THIRD PERIOD:	3815	Student-item-day encounters
FIFTH PERIOD:	3005	Student-item-day encounters
SIXTH PERIOD:	2915	Student-item-day encounters
TOTAL:	17,365	Student-item-day encounters

It might be well to mention at this time that Mr. Welsh considered the fifth and sixth period his "highest performance" classes.** The reader will notice that these two classes also accumulated the lowest student-item-day ratio of the five participating classes in the experiment.

STUDENT AWARENESS OF THE STUDY

Since it was felt that any unusual effort to prepare and achieve by the students would be randomly distributed throughout the week and not necessarily accrued consistently on any given day, the students were made aware that some type of study was being conducted. They were not, however, informed specifically regarding the nature of the experiment, but rather they were told that the study related to two areas:

---

*Hereafter referred to as: Student-item-days.

** No substantiating data is included in this report to support this statement.

1. Improved student teacher performance.
2. An attempt to allow them (the students) to improve grade levels by repeated short item quizzes on small areas of study.[8]

---

[8] R. E. Gadske, "A Comparison of two Methods of Teaching First Year High School Algebra." Sch. Sci. Math, pp. 635-40. 1933.

PROCEDURES AND TECHNIQUES

THE EXPERIMENTAL TREATMENT

For the purposes of this study, and in order that internal validity might be strengthened, it was deemed advisable to observe the following procedural techniques which were kept constant, whenever and wherever possible, throughout the experimental treatment.

1. The experimental treatments were administered each day at the beginning of the period.

2. The time allowed for test completion was kept constant within ten to fifteen minutes — approximating a mean of twelve minutes. This time included:

    a. Reading the question.
    b. Writing the answer.
    c. Necessary grading and paper collecting activities.

3. Efforts were made to keep the climate as relaxed and informal as possible.

4. The tests were consistently constructed, scored and recorded by a member of the research team OTHER than the one presenting the subject content to the class.

5. Assignments for outside of class study were made daily—consisting whenever and wherever possible of three pages of text material for each day's study load. This "three page" goal was usually realized. Each three page assignment was cumulative through the chapter, i.e.: Monday three pages, Tuesday three pages, Wednesday three pages with testing as follows: Monday three pages, Tuesday six pages, Wednesday nine pages and so on. This is explained more fully by the following chart:

CLASS DAY OF WEEK	DAILY ASSIGNMENT	MATERIAL COVERED
Monday	Three pages	Three pages
Tuesday	Three pages	Six pages
Wednesday	Three pages	Nine pages
Thursday	Three pages	Twelve pages
Friday	Three pages	Fifteen pages

6. The three page approach as used throughout resulted in a "staggering" of the chapter beginning and ending days. No effort was made either to accentuate or dampen this effect. An early consideration was to intentionally stagger the schedule to avoid the pile up of chapter iniating activities on Monday and the chapter terminating activities on Friday. However, after consultation with Dr. Thomsen it was decided to ignore this effect to secure a smooth flow of information over the testing areas. Upon adoption of the three page rule all assignments were subjected to identical treatment and

no emphasis was placed at any point. On some occasions an assignment could include the last page or two of one chapter and the first page or so of the following — whatever was needed to fulfill the three page requirement.

7. Each day Mr. Welsh was supplied with seven or more items on individual item cards. These cards were added to the previous day's cards (for each chapter). The cards were shuffled well and five cards picked at random by some selected student. This was repeated period by period for all five classes of General Science each testing day.

8. Item numbering was consecutive by the chapters. When questions were drawn randomly the number accompanied the question on the test. A five item test answer sheet might be numbered thus:

5. ____________
11. ____________
9. ____________
1. ____________

THREATS TO VALIDITY IN THE EXPERIMENT

Based on concepts from Campbell and Stanley, the following outline summarizes the major sources of invalidity as they pertain to this study.[9]

A. Internal threats:

1. History:

   Routine absenteeism presented a relatively small threat to this study so long as it remained more or less on a random basis. During, and because of, the recent civil disturbances in this city, however, absenteeism was no longer random, but related to a specific cause. More important—its occurance concentrated on two critical days in the study— Monday and Friday. This posed a real threat to the study: one that the researchers were powerless to control.

2. Maturation:

   No specific control applied—but because of the relatively short duration of the study the researchers feel that there was a minimal cause for invalidity from this source.

3. Testing:

   Because there was no pretest data taken, the threat of testing as well as testing—treatment interaction is not a relevant threat—no interaction can occur.[10] Additional control was attempted in this area through:

   (a) Random selection of individual five inch by seven inch test item cards.

   (b) Consecutive three page content assignment, daily, in the regular course text book.

4. Instrumentation:

   This was an area of general weakness and needs more consideration and attention in future studies. Lack of standardization was a continuing threat to the study.

---

[9] Donald T. Campbell and Julian C. Stanley, Experimental And Quasi—Experimental Designs For Research, Chicago: Rand McNally & Company, pp.5-6, 40-41. 1966.

[10] Experimental Design: Paradigms and Procedures (a class handout), pp. 11-13.

4. Instrumentation:

This was an area of general weakness and needs more consideration and attention in future studies. Lack of standardization was a continuing threat to the study.

5. Regression:

This source was not a threat to this study since there was no pretest data taken and the groups were not selected on the basis of their extreme scores. The use of mean differences throughout the study also offers a built in control in this area.

6. Selection bias:

There is the possibility of some degree of selection bias in the study because no effort was made to assign groups in a random manner, but rather to work within the total frame work of all five available ninth grade science classes that had been previously assigned for several weeks prior to the initiation of this experiment.

7. Mortality:

There was little threat to the study from drop-outs or out-transfers---only six such cases over the seven weeks period. However, the situations present on which only partial data was available were more numerous. The experimentors more or less "Swept this under the rug," as per Campbell and Stanley.[11]

8. Selection---maturation interaction:

As noted by chart on page 75 this effect seemed not to present a major threat to the study since all groups seemed to conform to a very similar relationship with one another at the initiation and at the termination of the experimental treatment, even though there was no randomization of subjects.

B. External:

1. Testing and X-interaction:

Some "sensitizing" effect possibly occured during the early days of the study but since any particular day should receive a random chance of benefit in this area the threat posed to validity was relatively small, and the control was based on a daily test at the same general time.

---

[11] Campbell and Stanley, op. cit., p. 15.

2. Selection and X-interaction:

Groups were not randomly selected from the schools on a city-wide basis and Bellevue Junior High School is not equivalent to some of the "better" schools in the system because of a recent-years deterioration in the socio-economic background of the student body. Because of this, there may be a selection and X-interaction threat to the generalizability of the study.

3. Reactive arrangements:

These investigators felt that any efforts on the part of the students to alter performance as a result of situational knowledge would be randomly distributed over the Monday through Friday weekly testing span. For this reason the researchers did not consider that this source of invalidity posed a major threat to the study.

4. Multiple X-interference:

The researchers feel that this threat was under control because there was a one treatment per day design used throughout the experiment.

5-6. Irrelevant responses and irrelevant replicability did not offer any threats to the study at the time and under the conditions of its instrumentation.

ANOTHER SOURCE FOR INTERACTION

Interaction, or the possible effect of each cell on its neighboring cell or cells, in the design described herein is rather high, over all, for this experimental study. One particular possibility not mentioned above that continues to disturb the investigators relates to the manner in which the individual test items were selected.

As described elsewhere in this report, the test items for each day accumulated throughout the five test day week, and for each test day five items were randomly selected for that particular day's experimental treatment. Let the reader again refer to a chart:

	MON.	TUES.	WED.	THURS.	FRI.
NEW TEST ITEMS ADDED	6 to 10	6 to 10	6 to 10	6 to 10	6 to 10
PREVIOUS TEST ITEMS ADDED	None	All 6-10 Items from Monday.	All Mon. + Tues. items.	All Mon.+ Tues.+Wed. items.	All Mon. + Tues.+ Wed.+ Thurs. items.
TEST ITEMS RANDOMLY SELECTED FOR X	5	5	5	5	5

TOTAL: 25 items selected from 30 to 50 submitted weekly.

During the second or third week of the study the researchers began to question some possible source for invalidity posed by interaction at this level. Control was attempted by random selection of all items as indicated but the investigators feel that a tighter control is desirable for any future experimental effort.

SUMMARIZING ACCORDING TO CAMPBELL AND STANLEY[12]

INTERNAL SOURCES:

-	History
+	Maturation
+	Testing
-	Instrumentation
+	Regression
?-	Selection
?+	Mortality
+	Interaction of selection and maturation

EXTERNAL SOURCES:

+	Testing X interaction
?-	Selection X interaction
+	Reactive arrangements
+	Multiple X interaction

---

[12] Ibid., p. 56.

OTHER PROBLEMS IN THE STUDY

It was to be expected that some students would tend to study towards a particular type of examination. In an effort to control this particular threat to validity, inherent in the experimental treatment, the types of items and material emphasis was shifted throughout the study.

Another threat to validity, that was anticipated in this study related to spontaneous improvement in scores resulting from students becoming test wise and experienced. Efforts to control this threat to validity included a gradual increase in item difficulty as the study progressed.

One problem not anticipated by the investigators was related to cribbing or cheating by a small number of the students. The evidence of these deceptive activities was discovered by the researchers upon initiation of the statistical treatment on the data. All test and scoring calculations were routinely rechecked and corrected wherever discrepancies were discovered. Nevertheless this remains one of the major threats to the validity of the study; any future work along these lines will demand more attention to this problem and more efforts to correct the inadequacy.

Several interruptions related to city-wide civil disturbances and intervening school activities prevailed and resulted in limited or restricted research activities for these dates thereby reducing available data to the study. On April 3rd the fifth and sixth period classes received no experimental treatment, and on April 29th the fifth period was again omitted from the testing procedure.

No experimental treatment was carried out on the following dates:

March 28, 1968
March 29, 1968
April 5, 1968
April 8, 1968
April 12, 1968
April 15, 1968

Experimental subject losses related to student out-transfers during the course of this study amounted to six.

Student absences totaled three hundred five with the following distribution among the classes:

PERIOD	ABSENCES
First Period	72
Second Period	55
Third Period	68
Fifth Period	49
Sixth Period	61
Total	305

As indicated by the table on page 23 of this report, student subjects present and available for the experimental treatment during the seven weeks course of this study amounted to, by classes:

CLASS PERIOD	NUMBER OF STUDENTS TESTED
First Period	740
Second Period	786
Third Period	763
Fifth Period	601
Sixth Period	583
Total	3,473

Time was a crucial factor in carrying out the experimental treatments of this study and the investigators feel that a complete school year should be devoted to any future data gathering activities of a related nature.

OTHER RECOMMENDATIONS FOR IMPROVING THE STUDY

In regard to cribbing or cheating on the test one alternative would be for the researchers to score the test themselves—a rather time consuming activity. Another approach to solving the problem would be to furnish preprinted multiple item score cards and hand punchers. The students would mark only the correct choice by punching the indicated mark for a selection. This approach could be adapted to automated procedures of data processing and computer scoring and grading. They could also take the examination in ink, allowing no erasure or changes. Still another approach would be to furnish red ink or pencil for all scoring activities by the students. Another possibility would be to recruit volunteer student assistants whose sole responsibility would be the grading of the papers. The latter would probably be the method of choice at this time. The investigators realize that the above measures are no panacea but do feel that they would effect some reduction in the incidence of cheating on the test and therefore enhance test validity and reliability.

STATISTICAL ANALYSIS—RESULTS AND CONCLUSIONS

The statistical analysis design chosen for this research report was intended to render a day by day analysis of variability for a day versus day comparison over each of five days by five days horizontally and vertically, determining F values for each cell as indicated by the following row/column table:

	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
MON		Monday versus Tuesday F	Monday versus Wednesday F	Monday versus Thursday F	Monday versus Friday F
TUES	Tuesday versus Monday F		Tuesday versus Wednesday F	Tuesday versus Thursday F	Tuesday versus Friday F
WED	Wednesday versus Monday F	Wednesday versus Tuesday F		Wednesday versus Thursday F	Wednesday versus Friday F
THUR	Thursday versus Monday F	Thursday versus Tuesday F	Thursday versus Wednesday F		Thursday versus Friday F
FRI	Friday versus Monday F	Friday versus Tuesday F	Friday versus Wednesday F	Friday versus Thursday F	

This method utilized the F test which compares the variability within samples to the variability among two or more samples to determine the probability that the variance among samples is due to chance and not due to sampling from populations having different means. In other words, the techniques as used in this study were designed to determine whether the groups diverged sufficiently to be considered as no longer representative of the same pupulation. The values of F were determined by the use of appropriate tables and reported accordingly in charts and tables included in this report.

More specifically, Monday was subjected to an analysis of variance comparison with Tuesday and an F ratio determined and compared with F table values. Monday was then compared with Wednesday and so on through Friday. Each of the days in the week were similarly compared with every other day in the week. The computed F values were then compared with table F values at the one percent (.01) and the five percent (.05) levels of confidence to make the necessary determinations regarding significance of the data obtained.

FURTHER ANALYSIS WITH THE CHARTS

As indicated by summary chart number one the weekly week-day variations held to a rather similar pattern of fluctuation with the one exception of Friday (the short black line). Friday, however, represented a very short term situation as regards the experimental treatment since there were only four Friday's that received the experimental treatment during the entire seven weeks of the study. Nevertheless, there was no indication of any significant differences between any two given days, according to the analysis of variance done on the data obtained during the seven weeks period.

The bar graph charts indicate the standing of the classes based on the over all means. There are daily charts plus a summary chart for all five week days.

The sum of the means charts indicate a slight superiority for the fifth period class with a grand mean of 3.57 and the sixth period group follows closely with a grand mean of 3.36. These performances were expected by the investigators.

The expected low in the first period, however, did not occur. This dubious honor was taken by the third period class with a 2.93, while the first period scored a 3.06.

The expectations mentioned above were based on experience and classroom records recorded by Mr. Welsh and L. Joyner (as a student teacher) prior to the initiation of any experimental treatment related to this research effort.

The performance standing for any class is a chance occurance in the respect that no ability grouping is attempted at Bellevue Junior High School.

MEANING OF THE ANALYSIS

The investigators evaluated the results of the experiment by subjecting the data to an analysis of variance wherein the between-sample variance was compared with the within-sample variance in an ANOVA table. The between sample variance is a measure of the treatment effect contaminated by error and is obtained when only the average means being compared are used as the data. The within sample variance is a measure or estimate of the error which is responsible for the above contamination and is an over all variance computed for all the data being compared. The within variance also indicates the size of the errors, which in being randomly distributed over the two samples, may actually have produced the between-sample variance. The mean square ratio of one variance is compared with the other to obtain a computed F which is then analyzed according to predetermined table values for this comparison.[13]

As indicated by the series of ANOVA tables included in this report, and other tables and charts detailing the statistical findings--no significant difference in the means was found at the confidence levels now indicated. The investigators therefore failed to reject null at these levels for any data included and analyzed in this report. It is therefore concluded that NO SIGNIFICANT RELATIONSHIP EXIST BETWEEN ANY PARTICULAR DAY OF THE WEEK AND STUDENT ACHIEVEMENT AS MEASURED BY DAILY TEST ADMINISTRATION.

PROBLEMS WITH THE ANALYSIS

According to Doctors Thomsen and Fortune, the analysis as reported herein is subject to a geometric pile up of probability errors, horizontally by days, and "stacking" of this error, vertically by weeks.[14]

This major problem as confronted in the analysis of the data obtained in the study related to, and was primarily precipitated by, the incomplete data cells in the analysis of variance design. This element occuring in the study precluded the simple analysis of variance approach as permitted when all cells are filled.

The statistical design, basically an analysis of variance with incomplete blocks by days and nested intervals by weeks, was subjected to a modification which permitted the inclusion and comparison of only those days on which experimental treatment was carried out and data was obtained. Since Doctors Thomsen and Fortune, as mentioned above, suggested that this is primarily an approximation and therefore vulnerable to a pile-up of probability errors as stated, the determined ratios might be defined as F ' rather than F.[15] These workers suggest that any future analysis of these data include a chi squares or Latin squares analysis and an ANOVA based on the incomplete block, nested interval techniques as described by Scheffé.[16]

This writer plans to submit such an analysis as a supplementary inclusion to this report, if time permits.

---

[14] Dr. Donald Rickhart Thomsen, Associate Professor of Education, Memphis State University, Memphis, Tennessee and Dr. Jim Carleton Fortune, Assistant Professor of Education also at Memphis State University.

[15] Henry Scheffé, The Analysis of Variance, New York: John Wiley & Sons, Inc., pp. 147-88. 1959.

[16] Ibid., pp.413-15.

CONCLUSION

A brief look at the enclosed charts and tables will reveal that these investigators failed to reject null at the one percent (.01), or the five percent (.05), confidence levels for all comparisons considered.

Obviously the apparent treatment effects were small when compared with the errors which could have produced them. Since it is a fundamental principle in the analysis of variance that the greater the variability among the samples in the study, the larger will be the errors to which the experimental comparison is liable and the more suspect will be the observed differences between the samples, then it must follow, therefore, that chance alone could have produced differences as large as, or larger than, the differences observed in this study. It must be concluded then that the data obtained and analyzed for the purposes of this study failed to reveal any significant relationships, or effects, related to the particular week day of test administration on ninth grade student achievement as measured by daily testing over a seven week period at Bellevue Junior High School.

RECOMMENDATIONS FOR FUTURE STUDIES IN THIS AREA

These workers are of the opinion that more investigation should be conducted in the area of this study. They feel that the results are inconclusive and studies are recommended which would obtain more data under conditions of tighter controls. In this regard the following recommended which would obtain more data under conditions of tighter controls. In this regard the following recommendations for future studies are set forth:

1. A minimum period of one full school year should be devoted to the experimental treatment and data gathering activities. The seven weeks period of this study was too susceptible to class time lost and student absent rates; a longer period of time would insure more complete data for analysis.

2. Some effort should be made to randomize the samples for future study. This would permit a period by period evaluation and further reduce many internal threats.

3. If at all possible some standardized testing should be available for the selection of test items. These investigators were confronted with situations that may have reflected item or test difficulty rather than a week day influence--more effort needs to be made to separate these two effects.

4. Increased effort is needed to standardize the technique of the experimental treatment as well as the testing atmosphere. There is some question in the minds of the researchers regarding the conscientiousness of student attitudes towards these testing situations.

5. These investigators also feel that consideration should be given to separation of record grading and data recording techniques as well as discontinuance of the procedure for recording student names on each test. It is felt that this would help to further reduce test anxiety and possibly lessen the incidence of cribbing and cheating on the test by the students.

6. If politically feasible, acquisition of individual student I.Q. and previous achievement scores is to be desired in future efforts.

7. Another interesting prospect for future determinations by daily testing is related to applications of the experimental treatment to the evaluation of achievement by student subjects in elective areas rather than required courses.

8. The researchers also feel that some effort to obtain a study-habits-profile on the participating student subjects would provide some interesting material relevant to the study.

9. Related to number one above — effort should be made to include simultaneous replicability at several randomly selected schools in the city as well as over the suburbal and rural areas of West Tennessee.

10. It is also recommended that any future studies in this area be designed to incorporate electronic data processing equipment and technique into the statistical analysis.

11. Related to ten above — computer based testing and grading should be used when and where possible.

SUMMARY

In summary then, these workers are of the opinion that the data obtained, analyzed, evaluated and presented in this report are not conclusive and the study was confronted by too many variables and threats to validity which the researchers failed to adequately control. It is also the feeling of these investigators, however, that there has been established an adequate basis and indication for more extensive and precisely controlled work in this area.

PROJECT ON TEST DEVELOPMENT

January 3, 1968

by

Lucy Reap
Theresa Browning
Lynn Wells

INTRODUCTION:

All of the members of the Undergraduate Research Training Program at Memphis State University were assigned the project of developing a test and checking its reliability and validity; our team, composed of Theresa Browning, Lucy Reap, and Lynn Wells, wanted to work with something that we thought might be useful to us as future teachers. While preparing for our second semester research project, all of us had become interested in attitudinal scales and measurements. Since two of us are future mathematics teachers, we selected algebra as the school subject area to test and to use in correlation with our attitude scale.

With the aid of algebra textbooks checked out from the Memphis State Curriculum Lab, we developed a comprehensive algebra test. The attitude scale was designed in the form of Likert's attitudinal scales. Our attitude scale comprised six questions, three of which were distractors, with alternating favorable and unfavorable response choices.

We received permission from the principal of one of the city schools to administer our tests in his school; and, we enlisted the cooperation of a first-year algebra teacher, who offered us the students in her three classes as subjects.

PILOT TEST:

The pilot test was administered to the twenty-two students of the 9:45 a.m. class. We ran an item analysis on the fifteen items of the algebra test; results are shown on Table I.

The algebra test was divided into two parts by flipping a coin. The first part included items 2, 3b, 4, 5, 6, 7, 11, 15; the second part included items 1, 3a, 8, 9, 10, 12, 13, and 14. Using the split-half method of determining reliability, the reliability coefficient was established as .518347. (See Table II.)

This study's main point of interest was the determination of the effect that knowledge of results has in the development of the student's attitude toward the test itself. For this reason, the test was administered in two different ways. One half of the class received immediate feedback as to each student's score on the test. This portion of the class was given the answers to the algebra test, written on the attitude scale. The other portion of the class did not receive answers prior to responding to the attitude scale. The Pearson formula was used in the determination of the correlation between the algebra test and the attitude scale given to both groups. The correlation coefficients were as following: Immediate knowledge of the test results and the score on the attitude scale, -.612; no knowledge of the test results and the score on the attitude scale, .0154. Results are shown on Tables III a and III b.

FINAL TESTS:

From the results of the pilot test and suggestions by the supervising teacher, we increased the difficulty of the algebra test. The revised test was administered to two classes, an 8:45 a.m. class and a 2:15 p.m. class.

An item analysis of the fourteen items was made on the morning class of twenty-nine students; results are shown on Table IV a.

By flipping a coin, we divided the final test into two parts. The first part consisted of items 3a, 3b, 4, 5, 8, 9, 11, 12, 14; the second part contained items 1, 2, 6, 7, 10, 13. By use of the split-half method, the reliability coefficient of the test as given to the morning class was .7301. (See Table V.)

The students in the morning class did not know the correct answers to the algebra test before taking the attitude scale. The Spearman Correlation Coefficient between the algebra test and the attitude scale for this class was -.321. (See Table VI.)

The results of the item analysis made on the afternoon class of twenty students are shown on Table IV b.

By using the split-half method and the same random manner of item division, the reliability coefficient was found to be .5146. (See Table VII.)

Before taking the attitude scale, the students in the afternoon class were given the correct answers. For this class, the Spearman Correlation Coefficient was -.430. (See Table VIII.)

In combining the results of the morning class and the afternoon class, the over-all reliability of the tests was .6633. (See Table IX).

CONCLUSION:

Although Kerlinger states "High reliability is no guarantee of good scientific results ....", Downie says that on an ordinary class room test a reliability "of .50 might be typical." Our results exceed Downie's expectations of a classroom test.

On the final testing, the correlation coefficients show that the students who made the highest scores on the algebra tests rated the test relatively easy on the attitude scale. This was true for both classes taking the test — the one that had knowledge of the test scores prior to taking the attitude test and the one that had no knowledge of test performance before responding to the attitude scale. This would seem to indicate that immediate feedback of test results has no bearing in the student's attitude toward the test itself. However, the correlations obtained on the pilot test were slightly different than those obtained on the final tests. This could be due to the change in testing conditions. During the final tests, the class which was given the answers to the algebra test was completely separated from those who were not given the correct answers prior to responding to the attitude scale. Therefore, further studies in this area would be interesting in the light of the contrasting results of the pilot test and the final tests.

TABLE I

PILOT TEST

ITEM ANALYSIS

	CORRECT RESPONSES	INCORRECT RESPONSES
1.	17	5
2.	19	3
3. (a)	16	6
(b)	15	7
4.	18	4
5.	12	10
6.	18	4
7.	22	0
8.	22	0
9.	13	9
10.	0	22
11.	8	14
12.	2	20
13.	18	4
14.	19	3
15.	21	1

TABLE II

RELIABILITY OF PILOT TEST

$$r_{1I} = \frac{2r_{\frac{1}{2}\frac{1}{1}I}}{1 + r_{\frac{1}{2}\frac{1}{I}I}} \qquad r_{\frac{1}{2}\frac{1}{1I}} = \frac{P}{x\ \ y}$$

$$P = \frac{(x\,y)}{N} - \frac{x}{N}\ \frac{y}{N}$$

$$\sigma x = \frac{(x^2)}{N} - \frac{x}{N}^{2}\ 2$$

$$\sigma y = \frac{(Y^2)}{N} - \frac{Y}{N}$$

X	Y	X-Y	X^2	Y^2
7	4	28	49	16
6	6	36	36	36
6	5	30	36	25
4	4	16	16	16
5	3	15	25	9
7	7	49	49	49
8	4	32	64	16
8	5	40	64	25
7	6	42	49	36
6	5	30	36	25
7	6	42	49	36
6	3	18	36	9
4	4	16	16	16
5	3	15	25	9
5	5	25	25	25
7	5	35	49	25
7	5	35	49	25
5	5	25	25	25
5	3	15	25	9
5	6	30	25	36
6	6	36	36	36
6	6	36	36	36
Σ132	Σ106	Σ646	Σ820	Σ540

$$P = \frac{646}{22} - \left(\frac{132}{22}\right)\left(\frac{106}{22}\right) = 29.3636 - (6)\ (4.8181) = 29.3636 - 28.9086 = .4550$$

$$\sigma x = \sqrt{\frac{820}{22} - \left(\frac{132}{22}\right)^2} = \sqrt{37.2727 - 36} = \sqrt{1.2727} = 1.128$$

$$\sigma y = \sqrt{\frac{540}{22} - \left(\frac{106}{22}\right)^2} = \sqrt{24.5454 - 23.2141} = \sqrt{1.3313} = 1.153$$

$$\sigma x\ \sigma y = (1.128)(1.153) = 1.300584$$

$$\frac{P}{\sigma x\ \sigma y} = \frac{.4544}{1.300584} = .349843$$

$$\frac{2 \times .349843}{1 + .349843} = \frac{.699686}{1.349843} = r_{1I} = .518347$$

TABLE III a

ALGEBRA

SCORE ATTITUDE

X	Y	$X-\bar{X} = x$	$Y-\bar{Y} = y$	x-y	x^2	y^2
11.5	4.74	1.68	-2.54	- 4.2672	2.8224	6.4516
7.5	10.98	-2.32	3.70	- 8.5840	5.3824	13.6900
8.0	15.00	-1.82	8.72	-15.8740	3.3124	76.0384
11.5	5.49	1.68	-1.79	- 3.0072	2.8224	3.2041
10.5	7.26	.68	.02	- .0136	.4624	.0004
9.0	4.50	- .82	-2.78	2.2796	.6724	7.7284
7.0	7.74	-2.82	.46	- 1.2972	7.9524	.2116
11.0	5.40	1.18	-1.88	- 2.2184	1.3924	3.5344
9.5	10.25	- .32	2.97	- .9504	.1024	8.8209
10.0	3.75	.18	-3.53	- .6354	.0324	12.4609
12.5	5.01	2.68	-2.27	- 6.0836	7.1824	5.1529
Σ108.0	Σ80.12			Σ-40.6478	Σ32.1364	Σ137.2936

$\bar{x} = 9.82 \quad \bar{Y} = 7.28$

$$SD_x = \sqrt{\frac{32.1364}{11}} = \sqrt{2.9215} = 1.71$$

$$SD_y = \sqrt{\frac{137.2936}{11}} = \sqrt{12.4812} = 3.53$$

$$r = \frac{-40.6478}{11(1.71)(3.53)}$$

$$r = \frac{-40.6478}{66.3993}$$

$$r = -.612$$

TABLE III b

ALGEBRA

Score	Attitude	$X-\bar{X} = x$	$Y-\bar{Y} = y$	$x \cdot y$	x^2	y^2
10.0	4.98	-.59	-.49	.2891	100	24.8004
10.5	7.23	-.09	1.76	-.1584	110.25	52.2729
13.0	6.24	2.41	.77	1.8557	169	38.9376
12.5	4.50	1.91	-.97	-1.8527	156.25	59.9076
12.5	7.74	1.91	2.27	4.3357	156.25	59.9076
12.0	5.49	1.41	.02	.0282	144	30.1401
7.5	3.00	-3.09	-2.47	7.6323	56.25	9.0000
9.0	3.51	-1.59	-1.96	3.1164	81	12.3201
11.01	4.74	.41	-.73	-.2993	121	22.4676
7.5	6.99	-3.09	1.52	-4.6968	56.25	48.8601
11.0	5.76	.41	.29	.1189	121.00	33.1776
Σ116.5	Σ60.18			Σ10.3691	Σ1271.25	Σ352.1340

$$SD_x = \sqrt{\frac{1271.25}{11}} = \sqrt{115.5681} = 10.75$$

$$SD_y = \sqrt{\frac{352.1340}{11}} = \sqrt{32.0121} = 5.66$$

$$r = \frac{10.3691}{11(10.75)(5.66)}$$

$$r = \frac{10.3691}{669.2950}$$

$$r = .0154$$

TABLE IV a

Item Analysis A.M. Class

	Correct Responses	Incorrect Responses
1.	15	14
2.	22	7
3.	(a) 25 (b) 12	(a) 14 (b) 17
4.	23	6
5.	10	19
6.	15	14
7.	19	10
8.	0	29
9.	5	24
10.	6	23
11.	8	21
12.	10	19
13.	11	18
14.	6	23

TABLE IV b

Item Analysis P.M. Class

	Correct Responses	Incorrect Responses
1.	14	6
2.	13	7
3.	(a) 15 (b) 13	(a) 5 (b) 7
4.	20	0
5.	7	13
6.	11	9
7.	14	6
8.	2	18
9.	0	20
10.	8	12
11.	2	18
12.	6	14
13.	7	13
14.	4	16

TABLE V

Reliability A.M. Class

	Scores On 1st Part of Test x	Scores On 2nd Part of Test y	xy	x^2	y^2
	1	3	3	1	9
	2	4	8	4	16
	4	7	28	16	49
	3	2	6	9	4
	1	2	2	1	4
	3	4	12	9	16
	6	8	48	36	64
	3	2	6	9	4
	3	3	9	9	9
	0	1	0	0	1
	2	3	6	4	9
	2	4	8	4	16
	1	2	2	1	4
	2	2	4	4	4
	6	3	18	36	9
	4	3	12	16	9
	3	1	3	9	1
	4	3	12	16	9
	3	5	15	9	25
	3	3	9	9	9
	4	4	16	16	16
	1	2	2	1	4
	3	1	3	9	1
	4	6	24	16	36
	2	3	6	4	9
	5	5	25	25	25
	3	5	15	9	25
	3	3	9	9	9
	3	2	6	9	4
Totals	84	96	317	300	400

$$p = \frac{\Sigma(xy)}{N} - \left(\frac{\Sigma(x)}{N}\right)\left(\frac{\Sigma(y)}{N}\right)$$

$$p = \frac{317}{29} - \left(\frac{84}{29}\right)\left(\frac{96}{29}\right)$$

$$p = 1.3424$$

$$\sigma x = \sqrt{\frac{\Sigma(x^2)}{N} - \left(\frac{\Sigma x}{N}\right)^2}$$

$$\sigma x = \sqrt{\frac{300}{29} - \left(\frac{84}{29}\right)^2}$$

$$\sigma x = 1.39$$

$$\sigma y = \sqrt{\frac{\Sigma(y^2)}{N} - \left(\frac{\Sigma y}{N}\right)^2}$$

$$\sigma y = \sqrt{\frac{400}{29} - \left(\frac{96}{29}\right)^2}$$

$$\sigma y = 1.68$$

$$r = \frac{P}{\sigma x \; \sigma y} \quad \text{(correlation)}$$

$$r = \frac{1.3424}{(1.39)(1.68)}$$

$$r = .5749 \qquad r = r_{\frac{1}{2}\frac{1}{1\,I}}$$

$$r = \frac{2r_{\frac{1}{2}\frac{1}{1\,I}}}{1 + r_{\frac{1}{2}\frac{1}{1\,I}}}$$

$$r = \frac{2\,(.5749)}{1 + .5749}$$

$$r = .7301 \text{ (reliability}$$

TABLE VI

Correlation A.M. Class

X Alg. Test	Y Att. Test	Rank X	Rank Y	Diff.	$(\text{Diff})^2$
3.5	3.50	25.0	4.0	21.0	441.00
5.0	2.67	16.0	15.0	1.0	1.00
10.0	1.58	2.0	28.0	26.0	676.00
5.0	2.58	16.0	16.5	.5	272.25
3.5	2.58	25.0	16.5	8.5	272.25
6.5	3.58	10.5	3.0	7.5	56.25
13.0	1.42	1.0	29.0	28.0	784.00
4.5	3.00	20.5	9.0	11.5	132.25
5.5	3.00	12.5	9.0	3.5	12.25
.5	3.17	29.0	7.0	22.0	484.00
4.5	2.83	20.5	11.0	9.5	90.25
5.0	2.33	16.0	20.0	4.0	16.00
2.5	2.50	27.5	18.0	9.5	90.25
3.5	3.00	25.0	9.0	16.0	256.00
7.5	2.33	5.5	20.0	14.5	210.25
6.5	2.75	10.5	13.0	2.5	6.25
4.5	1.67	20.5	27.0	6.5	42.25
7.0	2.25	8.0	22.0	14.0	196.00
7.0	3.67	8.0	2.0	6.0	36.00
5.0	2.75	16.0	13.0	3.0	9.00
7.0	3.33	8.0	6.0	2.0	4.00
2.5	4.00	27.5	1.0	26.5	702.25
4.0	2.08	23.0	24.0	1.0	1.00
9.5	2.33	3.0	20.0	17.0	289.00
4.5	1.83	20.5	26.0	5.5	30.25
9.0	3.42	4.0	5.0	1.0	1.00
7.5	2.75	5.5	13.0	7.5	56.25
5.5	2.08	12.5	24.0	11.5	132.25
5.0	2.08	16.0	24.0	8.0	64.00
Total					5363.50

Note: This class did not know the correct answers on the algebra test before taking the attitude test.

$$\text{rho} = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

$$\text{rho} = 1 - \frac{6(5363.50)}{29(29^2-1)}$$

$$\text{rho} = -.321$$

Interpretation:
Students who made high on the algebra test rated the test as relatively easy.

TABLE VII

Reliability P.M. Class

	Scores On 1st Part of Test x	Scores On 2nd Part of Test y	xy	x^2	y^2
	5	6	30	25	36
	5	5	25	25	25
	4	5	20	16	25
	4	4	16	16	16
	3	1	3	9	1
	3	3	9	9	9
	2	3	6	4	9
	1	3	3	1	9
	4	5	20	16	25
	2	4	8	4	16
	5	2	10	25	4
	1	4	4	1	16
	3	3	9	9	9
	4	3	12	16	9
	4	4	16	16	16
	3	4	12	9	16
	4	5	20	16	25
	5	3	15	25	9
	2	1	2	4	1
	3	1	3	9	1
Totals	67	69	243	255	277

$$p = \frac{\Sigma xy}{N} - \left(\frac{x}{N}\right)\left(\frac{y}{N}\right)$$

$$p = \frac{243}{20} - \left(\frac{67}{20}\right)\left(\frac{69}{20}\right)$$

$$p = .5925$$

$$\sigma x = \sqrt{\frac{\Sigma x^2}{N} - \left(\frac{\Sigma x}{N}\right)^2}$$

$$\sigma x = \sqrt{\frac{255}{20} - \left(\frac{67}{20}\right)^2}$$

$$\sigma x = 1.23$$

$$\sigma y = \sqrt{\frac{\Sigma y^2}{N} - \left(\frac{\Sigma y}{N}\right)^2}$$

$$\sigma y = \sqrt{\frac{277}{20} - \left(\frac{69}{20}\right)^2}$$

$$\sigma y = 1.39$$

$$r = \frac{P}{\sigma x \ \sigma y}$$

$$r = \frac{.5925}{(1.23)(1.39)}$$

$$r = .3465 \text{ (correlation)}$$

$$r = 2\, r_{\frac{1}{2}} \frac{1}{1\ I}$$

$$r = \frac{2(.3465)}{1.3465}$$

$$r = .5146 \text{ (reliability)}$$

TABLE VIII

Correlation   P.M. Class

X Alg. Test	Y Att. Test	Rank X	Rank Y	Diff.	$(\text{Diff.})^2$
10.0	1.90	1.0	20.0	19.0	361.00
9.5	3.25	2.0	7.5	5.5	30.25
8.5	1.92	3.0	19.0	16.0	256.00
7.0	3.67	9.5	5.0	4.5	20.25
4.0	2.33	17.5	15.5	2.0	4.00
5.0	2.83	14.0	11.5	2.5	6.25
4.0	2.83	17.5	11.5	6.0	36.00
3.5	4.17	19.0	1	18.0	324.00
8.0	2.00	4.5	17.5	13.0	169.00
5.0	3.42	14.0	6.0	8.0	64.00
7.0	3.83	9.5	3.5	6.0	36.00
5.0	3.25	14.0	7.5	6.5	42.25
6.0	4.08	12.0	2.0	10.0	100.00
7.0	2.00	9.5	17.5	8.0	64.00
7.5	2.50	6.5	14.0	7.5	56.25
7.0	2.92	9.5	10.0	0.5	.25
8.0	2.33	4.5	15.5	11.0	121.00
7.5	3.08	6.5	9.0	2.5	6.25
3.0	2.75	20.0	13.0	7.0	49.00
4.5	3.83	16.0	3.5	12.5	156.25
Total					1902.00

Note: This class did know the correct answers on the algebra test before taking the attitude test.

$$\text{rho} = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

$$\text{rho} = 1 - \frac{6(1902.00)}{20\ (20^2-1)}$$

$$\text{rho} = -.430$$

Interpretation: Students who made high on the algebra test rated the test as comparatively easy.

TABLE IX

Reliability - Combined Classes

$\Sigma X = 84 + 67 = 151$ $\qquad$ $\Sigma X^2 = 300 + 255 = 555$

$\Sigma Y = 96 + 69 = 165$ $\qquad$ $\Sigma Y^2 = 400 + 277 = 677$

$\Sigma XY = 317 + 243 = 560$ $\qquad$ $N = 29 + 20 = 49$

$$p = \frac{\Sigma XY}{N} - \left(\frac{\Sigma X}{N}\right)\left(\frac{\Sigma Y}{N}\right)$$

$$p = \frac{560}{49} - \frac{151}{49}\ \frac{165}{49}$$

$$p = 1.0516$$

$$\sigma x = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2} \qquad \sigma y = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$$

$$\sigma x = \sqrt{\frac{555}{49} - \left(\frac{151}{49}\right)^2} \qquad \sigma y = \sqrt{\frac{677}{49} - \left(\frac{165}{49}\right)^2}$$

$$\sigma x = 1.35 \qquad \sigma y = 1.57$$

$$r = \frac{P}{\sigma x\ \sigma y}$$

$$r = \frac{1.0516}{(1.35)(1.57)}$$

$r = .4962$ (correlation)

$$r = \frac{2r}{1 + r}$$

$$r = \frac{.9924}{1 + .4962}$$

$r = .6633$ (reliability)

NAME ______________________________ DATE ____________________

PART I.

Simplify:

(1) $x^4 - x^2 + 3x^3 - 2X + x^4 + 6 - 2x^4 + x^2 - x - 1$ (2) $\frac{20xy^2}{-4y}$

(3) If x = -5 and y = -5, find the value of:
a. x^2y b. $-x^3$

Solve:

(4) $A = \frac{h(b+c)}{2}$ If b = 8, c = 18, h = 14.

(5) $C = \frac{5}{9}(F-32)$ for F when C=10

(6) $p + .12p = 280$

(7) $1.7x + 604 = 2.3x + 148$

PART II.

(8) Multiply the product of c + 5 and c - 8 by c - 2.

(9) Multiply $a^2 - 6a + 9$ by a - 1.

(10) Divide $Y^2 - X^2$ by Y - X.

(11) Solve for x and y: $3x + y = 5$
$2 - y = 9$

(12) On Monday, the temperature was 20° above zero; on Tuesday, the temperature dropped to 7° below zero; on Wednesday, the temperature was 6° above zero. What was the total change in temperature in degrees over the three-day period?

(13) The sum of two numbers is 39. Two times the smaller number plus two equals one half the larger number. What are the two numbers?

(14) A man has $2.00 in nickels and dimes. One half the number of nickels times six divided by two is your number of dimes. How many nickels and dimes does he have?

ANSWERS:

1. ______________________________
2. ______________________________
3. ______________________________
4. ______________________________
5. ______________________________
6. ______________________________
7. ______________________________
8. ______________________________
9. ______________________________
10. ______________________________
11. ______________________________
12. ______________________________
13. ______________________________
14. ______________________________

NAME ______________________________ DATE ____________________

Please answer each of these six questinns in the following manner:

EXAMPLE: Today the weather is

Cold _____ _____ __X__ _____ _____ Hot

Still _____ _____ _____ _____ __X__ Windy

If the temperature is in the high 50's (approximately an "average" temperature), put an X in the third blank (see above). If it is very windy, put an X in the fifth blank.

Please be honest in answering these questions. We want each individual's opinion. These questions will in no way affect your grade and they will be held confidential.

1. I would rate Algebra as:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

2. The problems on this test were:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

3. Part I was:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

4. Part II was:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

5. In general, the homework assignments for this class have been:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

6. The explanations by the teacher in class have been:

Easy _____ _____ _____ _____ _____ Difficult
Complex _____ _____ _____ _____ _____ Simple
Clear _____ _____ _____ _____ _____ Confusing
Vague _____ _____ _____ _____ _____ Well-explained

NAME ______________________________ DATE ____________________

PART I.

(1) $x^2 - 2x + 6 - x^2 - 1$

(2) $\frac{20x^2}{-5x}$

(3) If $x = -5$ and $y = -5$, find the value of:
a. xy^2 b. $-x^3$

Solve:

(4) $A = \frac{h(b + b')}{2}$ If $b = 8$, $b' = 18$, $h = 14$.

(5) $C = \frac{5}{9}(F-32)$ for F when $C = 100$.

(6) $p + .12p = 280$

(7) $1.3x + 604 = 2.7x + 142$

PART II.

(8) Multiply $c + 5$ by $c - 8$

(9) Multiply $a^2 - 6a + 9$ by $a - 1$.

(10) Divide $x^2 - 6x + 9$ by $x - 3$.

(11) Divide $x^2 - y^2$ by $x - y$.

(12) Solve for x and y: $3x + y = 5$
$x - y = 1$

(13) What is the difference in degrees between 6° above zero and 18° below zero?

(14) The sum of two numbers is 40 and the larger exceeds three times the smaller by 4. What are the numbers?

(15) A man has 2 more nickels than dimes. He has $1.15 in all. How many coins of each kind has he?

ANSWERS:

1. ______________________________
2. ______________________________
3. ______________________________
4. ______________________________
5. ______________________________
6. ______________________________
7. ______________________________
8. ______________________________
9. ______________________________
10. ______________________________
11. ______________________________
12. ______________________________
13. ______________________________
14. ______________________________
15. ______________________________

THE EFFECT OF THE PARTICULAR WEEK DAY OF TEST ADMINISTRATION ON NINTH GRADE STUDENT ACHIEVEMENT AS MEASURED BY DAILY TESTING[1]

SUPPLEMENT NUMBER ONE
TO THE REPORT

[1]Further Analysis

SUBMITTED BY:
LEA W. JOYNER

JUNE 1968

INTRODUCTION TO THE SUPPLEMENT

COMMENTARY

The reader desiring to further inform himself regarding the specific techniques of the modern approach to Analysis of Variance will find:

1. No shortage of materials on the subject written for the more mathematically sophisticated. Scheffe and Gage are two sources included in the bibliography to this report.

2. Materials written for the person without a fairly extensive background in mathmatics are not quite so prevalent nor easy to come by.

This writer considers himself to be in the latter category and was hard pressed for published information that permitted the necessary understanding without seeking a higher level or professional interpretation of the material. This opportunity is taken to suggest, therefore, that one of the major obstacles to acquiring more quality research in the field of education might well be related to some aspect of this particular problem. It is the opinion of this writer that much of the material could be simplified in both the explanation and in the examples of usage and demonstrated applications. There is much need for some effort to produce MORE STRAIGHTFORWARD, SIMPLIFIED AND TO THE POINT statistical materials that would allow persons with otherwise acceptable training and backgrounds to conduct adequate analysis on data, competently obtained in well designed research studies. To say that every educational researcher must be a mathematics major is analogous to saying that every competent housekeeper and cook should also be a graduate economist and chemist in order to adequately interpret the household budget and the family cookbook.

The feeling of this writer is that qualified mathematicians could, and should, produce some materials along the lines of the Monarch Series or the Barnes and Nobles Series.[1] Such materials should be adequate and sufficiently detailed in content and clarified in explanation for the "nonmathematical" researcher to use as a necessary tool for more complete and reliable analysis of data.[2]

---

[1] Reinhard Herink, College Level Statistics, New York: Monarch Press, Inc., 1965. Herbert Arkin and Raymond R. Colton, Statistical Methods, New York: Barnes & Noble, Inc. 1967.

[2] Ibid. There is a current publication by Monarch Press with a too brief discussion that would be ideal if expanded to cover all areas of two way, three way, and four way Analysis of Variation and Analysis of Covariation.

DEFINITIONS

For purposes of this study and the included evaluation of data analysis the following definitions and/or symbols will be used:

1. N = The number of items in the situation.
2. $\bar{N}$ = The number of items in the study.
3. $\sum$ = The Sum of.
4. $\bar{\sum}$ = The GRAND SUM.
5. X = SCORE(s).
6. $\bar{X}$ = The ARITHMETIC MEAN of scores.
7. $\bar{\bar{X}}$ = The ARITHMETIC GRAND MEAN of all scores (also GM).
8. i = Mathematical subnotational reference to days that varied as: 1, 2, 3, 4, 5.
9. j = Mathematical subnotational reference to classes that varies as: 1, 2, 3, 4, 5.
10. k = Mathematical subnotational reference to weeks that varied as: 1, 2, 3, 4, 5, 6, 7.
11. H_0 = The NULL Hypothesis.
12. ANOVA = Analysis of VARIANCE.
13. ANOCOVA = Analysis of COVARIANCE.
14. SS = SUMS of the SQUARES obtained by: $\sum(\bar{X} - \bar{\bar{X}})^2 = SS$.
15. DF = DEGREES OF FREEDOM --usually obtained by N-1 or a related combination of the situational N-1.
16. For purposes of this study *SIGNIFICANT DIFFERENCE OR SIGNIFICANT VARIATION* will be determined by the variance ratio test.
17. The *VARIANCE RATIO OR F TEST* is often referred to as Snedecor's F test, since Snedecor computed tables for the variance ratio distribution, and named the ratio F, in honor of R. A. Fisher. It depends mathematically on Fisher's Z distribution, an extremely general and fundamental distribution which includes the normal distribution, the chi square (x^2) distribution, and Student's T distribution as special cases.[3]

---

[3] M. J. Moroney, *Facts From Figures*, Baltimore: Pelican Books, 1956, p. 233.

18. For purposes of this study the combined effects of two or more variables as determined by ANOVA and F comparison are considered to be, and are referred to as: INTERACTIONS.

19. Single variables judged to be possible sources of observable variations as determined by statistical analysis of F comparison are considered to be, and are referred to as: MAIN EFFECTS.

ABOUT THE ANOVA[4]

Moroney enthusiastically, and this writer believes appropriately, terms the modern statistical methods for the analysis of variation and covariation:

> "Undoubtedly one of the most elegant, powerful, and useful techniques in modern statistical method..."

This technique of analysis allows the researcher to reduce the total variation in the data under consideration to the particular elements of main effects and/or any combination of these effects that might possibly contribute to, or be associated with, potential sources of variability, the importance of which the researcher may desire to study or assess.

This technique is analogous to subjecting an organic biopsy to microscopic study and detailed analysis following a gross examination of the particular organ of concern. These investigators have tried to do much the same with the data obtained in the study.

They first reported findings based on a more superficial analysis as described in the preceding report on the study. With this enclosed supplement to that report they now attempt to discover more detailed and specific trends and relationships that might possibly be revealed by the relative components of these data as obtained in the study.

The heartbeat of the analysis used herein depends upon computing the variance of the data as the mean square (MS) deviation of each of the items from the grand mean ($\bar{\bar{X}}$ or GM) of all the items in the data set under consideration. Since, according to Moroney, a small sample tends to underestimate the true variance of the parent population a better estimate is usually obtained by dividing the sums of the squares $\boxed{\sum (X - \bar{\bar{X}})^2 = SS}$ by the number of degrees of freedom $\boxed{DF = (N-1)}$ in the sample as: $\sum \frac{(X - \bar{\bar{X}})^2}{N-1}$.[5]

---

[5] George W. Snedecor, Statistical Methods, Ames, Iowa: Iowa State College Press, 1946, pp. 222-25.

This approach to the data analysis requires that the investigator must compute the sum of the squares (SS) and the degrees of freedom (DF) for each source of variability considered.

As indicated above, the number of degrees of freedom on which any particular SS is computed will be one less than the number of items on which a given calculation is made, or DF = N-1.

The next consideration in the ANOVA is based on the assumption that the samples will not have the same average values but should vary "between" any two given sample groups according to the difference reflected by the particular members "within" each sample group—ALL ELSE REMAINING CONSTANT. In the words of Moroney then:

> "...What we should expect in these circumstances is that the variation between sample averages should be commensurate with the population variance as indicated by the variation within the individual samples. If it should prove that the 'between sample variation' were significantly greater than the 'within sample variation', then we should suspect that the samples were not, in fact, drawn from the same population, but from populations whose average values differed, so that on top of the 'within population variation' there existed also a 'between population variation'."

As stated elsewhere in this Supplement To The Report the analysis of variance technique is a complex and detailed approach that, being more sensitive to the respective sources of variation, was used herein to account for the performances within groups and performances between groups and make the proper determinations based upon the ratio of one source of variation to the other. In other words, the investigator was concerned at this point with the variation among the classes/days/weeks and the combinations of these variations that would proportionally be out of step with the degree of variation found within any particular sample among the member elements therein—thus indicating some possible source and/or degree of significant variation resulting from differential treatment appearing in the experiment, either intentionally or unintentionally.

It is hoped, therefore, that proper utilization of the ANOVA has resulted in an adequate accounting for each group's score in terms of variation among individual scores (class) as well as providing some indication for sources of variation, if any, caused by or related to some aspect of the experimental manipulation. As shown by the ANOVA tables of calculations, an attempt was also made to further indicate the variance between groups as a function of a single variable and also as a function of some combination of possible interaction between these variables.

The detailed explanation of these computations and the rationale behind each usage is rather involved and no attempt will be made to present an in depth discussion at this point.

To facilitate interpretation and estimations based on results of the study, the data and all computations are presented, wherever practical, in chart of tabular form. It is hoped that this will simplify the task of evaluation for the reader.

SUMMARY

A Brief Review Of Procedure:

1. The basic measure of variance, as used in this study, is the mean squared deviation from the grand mean, or notationally: $(\bar{X} - \bar{\bar{X}})^2$.
2. Remember also that the investigator was concerned first with the total amount of variation in the results, regardless of the source.
3. The next step was to calculate the variance within each group and establish its ratio to the computed between-groups variance.
4. The latter variation was attributed to the differential treatment of the sample and as the total between groups variance it was compared to the previously calculated within groups variance (see 3 above) for some degree of significance.
5. The ratio established in 3 and 4 above is sometimes referred to as the SIGNIFICANCE RATIO and as such it was compared with an F table of previously established values.[6]

---

[6] Please see discussion of F table on page 10 of this report.

THE STATISTICAL ANALYSIS OF VARIANCE

For purposes of this study significance was determined by ratio comparison of the between sample variance/within sample variance estimates with a Fisher's significance table as reproduced from George W. Snedecor.[7] The resulting estimates were stated in terms of an F and symbolized as Table Values which were either: (1) greater than the computed value (>) or (2) less than the computed value (<).

This writer considers the footnoting for each table as adequate for purposes of clarification and will at this time merely call the readers attention to the following items.

First, the column headings are:

1. SOURCE, or the different experimental conditions.
2. DF or degrees of freedom.
3. MS or mean squares.
4. F value as computed.

An understanding by the reader of the first (source) and last (F) are all that is required to make an adequate interpretation of the tables.

Second, for purposes of this study the F ratio was in all instances determined at two levels of confidence (1) the .01 or 1% level of confidence and (2) the .05 or 5% confidence level. The probability indication is that any variation of the degree revealed by the data processing analysis included herein would be expected to occur by chance alone (1) one time in every one hundred (1/100) and (2) five times in every one hundred (5/100) respectively. The ability to reject the Null Hypothesis (H_0) is indicated for each particular ratio throughout the comparison tables.

The nonsignificant mean effects and/or interactions with the resulting failures to reject the Null Hypothesis are likewise indicated by the tables.

The statistical analysis now follows and begins with a summation chart for the total variance followed by daily sub-total calculations, ANOVAS for main effects, interaction and sumation charts for purposes of revealed significance.

---

[7] Snedecor, *loc*. *cit*.

TOTAL VARIANCE: ALL SOURCES

FOR ALL TREATMENT WEEKDAYS---
Mondays Through Fridays

ACCORDING TO:

$$VAR_{total} = \sum \left[X_{ijk} - (X\ldots)^2 \ldots \right]$$

$$VAR_{total} = \sum \left(VAR_{Mons.} + VAR_{Tues.} \ldots + VAR_{Fris.} \right)$$

Mondays subtotal = 13.5041
Tuesdays subtotal = 32.4132
Wednesdays subtotal = 18.3431
Thursdays subtotal = 13.5041
Fridays subtotal = 05.6796

83.4441 = GRAND TOTAL FOR DAYS MONDAYS THROUGH FRIDAYS

MONDAY

$(\text{Score} - \text{GM})^2 = (\text{MON})^2 = \text{VAR}_{\text{Monday},}$

$(0.42)^2 = 0.1764$
$(0.41)^2 = 0.1681$
$(0.42)^2 = 0.1764$
$(0.69)^2 = 0.4761$
$(0.72)^2 = 0.5184$
$(0.77)^2 = 0.5929$
$(1.14)^2 = 1.2996$
$(0.07)^2 = 0.0049$
$(1.12)^2 = 1.2544$
$(0.77)^2 = 0.5929$
$(1.85)^2 = 3.4225$
$(0.88)^2 = 0.7744$
$(1.46)^2 = 2.1316$
$(1.60)^2 = 2.5600$
$(1.14)^2 = 1.2996$
$(1.26)^2 = 1.5876$
$(0.80)^2 = 0.6400$
$(0.05)^2 = 0.0025$
$(1.55)^2 = 2.4025$
$(0.92)^2 = 0.8494$
$(0.09)^2 = 0.0081$
$(0.43)^2 = 0.1849$
$(0.58)^2 = 0.3364$
$(0.38)^2 = 0.1444$

Monday subtotal $= 13.5041$

TUESDAY

$(\text{Score} - \text{GM})^2 \quad = \quad (\text{TUE})^2 \quad = \quad \text{VAR}_{\text{Tuesday}},$

$(0.49)^2 = 0.2401$

$(0.18)^2 = 0.0324$

$(0.24)^2 = 0.0576$

$(1.11)^2 = 1.2321$

$(0.08)^2 = 0.0064$

$(1.22)^2 = 1.4884$

$(0.14)^2 = 0.0196$

$(0.85)^2 = 0.7225$

$(1.36)^2 = 1.8496$

$(1.05)^2 = 1.1025$

$(0.93)^2 = 0.8624$

$(0.68)^2 = 0.4624$

$(0.56)^2 = 0.3136$

$(0.97)^2 = 0.9409$

$(0.62)^2 = 0.3844$

$(0.18)^2 = 0.3424$

$(0.32)^2 = 0.1024$

$(0.90)^2 = 0.8100$

$(0.95)^2 = 0.9025$

$(0.78)^2 = 0.6084$

$(0.61)^2 = 0.3721$

$(0.04)^2 = 0.0016$

$(0.95)^2 = 0.9025$

$(0.87)^2 = 0.7569$

$(0.42)^2 = 0.1764$

$(1.53)^2 = 2.3409$

$(1.34)^2 = 1.7956$

$(1.14)^2 = 1.2996$

$(0.66)^2 = 0.4356$

$(1.26)^2 = 1.5876$

$(2.11)^2 = 4.4521$

$(1.18)^2 = 1.3924$

$(1.38)^2 = 1.9044$

$(1.68)^2 = 2.8224$

Tuesday subtotal $= 32.4132$

WEDNESDAY

$(\text{Score} - \text{GM})^2 = (\text{WED})^2 = \text{VAR}_{\text{Wednesday},}$

$(0.20)^2 = 0.0400$

$(1.36)^2 = 1.8496$

$(0.70)^2 = 0.4900$

$(1.19)^2 = 1.4161$

$(0.45)^2 = 0.2025$

$(0.25)^2 = 0.0625$

$(0.64)^2 = 0.4096$

$(0.57)^2 = 0.3249$

$(0.44)^2 = 0.1936$

$(0.32)^2 = 0.1024$

$(0.12)^2 = 0.0144$

$(1.44)^2 = 2.0736$

$(0.68)^2 = 0.4624$

$(0.82)^2 = 0.6724$

$(0.60)^2 = 0.3600$

$(0.21)^2 = 0.0441$

$(0.01)^2 = 0.0001$

$(1.37)^2 = 1.8769$

$(1.36)^2 = 1.8496$

$(1.01)^2 = 1.0201$

$(1.27)^2 = 1.6129$

$(0.47)^2 = 0.2209$

$(0.32)^2 = 0.1024$

$(0.52)^2 = 0.2704$

$(0.24)^2 = 0.0576$

$(0.95)^2 = 0.9025$

$(0.32)^2 = 0.1024$

$(0.33)^2 = 0.1089$

$(0.14)^2 = 0.0196$

$(0.47)^2 = 0.2209$

$(0.99)^2 = 0.9801$

$(0.51)^2 = 0.2601$

$(0.14)^2 = 0.0196$

Wednesday subtotal $= 18.3431$

THURSDAY

$(\text{Score} - \text{GM})^2 = (\text{THUR})^2 = \text{VAR}_{\text{Thursday}},$

$(0.89)^2 = 0.7921$

$(0.93)^2 = 0.8649$

$(0.78)^2 = 0.6084$

$(0.82)^2 = 0.6724$

$(0.58)^2 = 0.3364$

$(0.34)^2 = 0.1156$

$(0.82)^2 = 0.6724$

$(0.51)^2 = 0.2601$

$(0.86)^2 = 0.7396$

$(0.42)^2 = 0.1764$

$(0.67)^2 = 0.4489$

$(0.12)^2 = 0.0144$

$(0.62)^2 = 0.3844$

$(0.96)^2 = 0.9216$

$(0.62)^2 = 0.3844$

$(0.06)^2 = 0.0036$

$(0.61)^2 = 0.3721$

$(1.03)^2 = 1.0609$

$(0.98)^2 = 0.9604$

$(0.18)^2 = 0.0324$

$(0.34)^2 = 0.1156$

$(0.42)^2 = 0.1764$

$(1.01)^2 = 1.0201$

$(0.28)^2 = 0.0784$

$(0.13)^2 = 0.0169$

$(0.43)^2 = 0.1849$

$(0.54)^2 = 0.2916$

$(1.22)^2 = 1.4884$

$(0.52)^2 = 0.2704$

$(0.20)^2 = 0.0400$

Thursday subtotal = 13.5041

FRIDAY

$(\text{Score} - \text{GM})^2 = (\text{FRI})^2 = \text{VAR}_{\text{Friday}}$

$(0.11)^2 = 0.0121$

$(0.61)^2 = 0.3721$

$(0.61)^2 = 0.3721$

$(0.60)^2 = 0.3600$

$(0.14)^2 = 0.0196$

$(0.70)^2 = 0.4900$

$(0.22)^2 = 0.0484$

$(0.11)^2 = 0.0121$

$(0.68)^2 = 0.4624$

$(0.09)^2 = 0.8100$

$(0.01)^2 = 0.0001$

$(0.53)^2 = 0.2809$

$(0.45)^2 = 0.2025$

$(0.62)^2 = 0.3844$

$(0.42)^2 = 0.1764$

$(0.74)^2 = 0.5476$

$(0.11)^2 = 0.0121$

$(0.32)^2 = 0.1024$

$(0.12)^2 = 0.0144$

$(1.00)^2 = 1.0000$

Friday subtotal = 5.6796

(Total variance calculations on page 12)

Table Number I

MAIN EFFECT: DAYS

ANALYSIS OF VARIANCE— Computation For Sums Of Squares

SS_{days} ACCORDING TO:

$$\sum_{i=1}^{5} \frac{1}{jk} \quad (\bar{X}_{i} \,..\, - \bar{\bar{X}} \,...)^2$$

$$\left(\begin{array}{l}\text{Total number of}\\ \text{scores for each}\\ \text{day in study}\end{array}\right) \quad \left(\begin{array}{l}\text{Mean for}\\ \text{each class}\\ \text{all classes}\end{array}\right) \quad - \quad \left(\text{GM}\right)^2$$

$$24(3.13 - 3.18)^2 = 0.0600$$

$$34(3.17 - 3.18)^2 = 0.0034$$

$$33(3.17 - 3.18)^2 = 0.0033$$

$$30(3.34 - 3.18)^2 = 0.7680$$

$$20(3.11 - 3.18)^2 = 0.0980$$

$$0.9327 = SS_{days}$$

Table Number 2

MAIN EFFECT: WEEKS

ANALYSIS OF VARIANCE--
Computation For Sums Of Squares

SS_{weeks} ACCORDING TO:

$$\sum_{k=1}^{7} \frac{1}{i\ j} \left(\bar{X}_{..k} - \bar{\bar{X}}_{...}\right)^2$$

$$\left(\begin{matrix}\text{Number of} \\ \text{weeks/situations}\end{matrix}\right) \left(\begin{matrix}\text{Day} \\ \text{mean}\end{matrix} - \begin{matrix}\text{Grand} \\ \text{mean}\end{matrix}\right)^2$$

$$29(3.06 - 3.18)^2 = 0.4176$$

$$29(2.98 - 3.18)^2 = 1.1600$$

$$30(2.93 - 3.18)^2 = 1.8750$$

$$26(3.57 - 3.18)^2 = 3.9546$$

$$28(3.36 - 3.18)^2 = 0.9072$$

$$8.3144 = SS_{weeks}$$

Table Number 3

MAIN EFFECT: CLASSES

ANALYSIS OF VARIANCE—
Computation For Sums Of Squares

SS_{classes} ACCORDING TO:

$$\sum_{j=1}^{5} \frac{1}{jk} \quad (\bar{X}_{.j.} - \bar{\bar{X}}_{...})^2$$

$$\left(\begin{array}{l}\text{Number of} \\ \text{classes/situations}\end{array}\right) \left(\begin{array}{l}\text{Day} \\ \text{mean}\end{array} \quad - \quad \begin{array}{l}\text{Grand} \\ \text{mean}\end{array}\right)^2$$

$$25(3.06 - 3.18)^2 = 0.3600$$

$$25(2.98 - 3.18)^2 = 1.0000$$

$$25(2.93 - 3.18)^2 = 1.5625$$

$$25(3.57 - 3.18)^2 = 3.8025$$

$$25(3.36 - 3.18)^2 = \underline{0.8100}$$

$$7.5350 = SS_{\text{classes}}$$

Table Number 4

INTERACTION: DAYS BY CLASSES

ANALYSIS OF VARIANCE
Computation For Sums Of Squares

$SS_{d\ X\ c}$ ACCORDING TO:

$$\sum_{j=1}^{5}\sum_{k=1}^{7}(\bar{X}_{ij\cdot} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot} + \bar{\bar{X}}_{\cdot\cdot\cdot})^2$$

$$\left(\begin{array}{l}\text{No. of scores}\\ \text{in individual}\\ \text{cell}\end{array}\right)\left(\begin{array}{l}\text{Class mean for}\\ \underline{\text{all}}\text{ Mondays,}\\ \text{Tuesdays, Wednesdays}\\ \text{Thursdays and}\\ \text{Fridays}\end{array} - \begin{array}{l}\text{Mean for}\\ \text{each day}\\ \text{of the}\\ \text{study}\end{array} - \begin{array}{l}\text{Mean of}\\ \text{whole}\\ \text{group for}\\ \text{individual}\\ \text{day}\end{array} + \text{GM}\right)^2$$

MONDAY

$5(2.61 - 3.06 - 3.13 + 3.18)^2 = 0.8000$

$5(3.17 - 3.06 - 3.17 + 3.18)^2 = 0.0720$

$5(3.08 - 3.06 - 3.17 + 3.18)^2 = 0.0045$

$4(3.38 - 3.06 - 3.34 + 3.18)^2 = 0.1024$

$5(3.48 - 3.06 - 3.11 + 3.18)^2 = 1.2005$

$2.1794 = \text{Subtotal } SS_{\text{Mon.}}$

TUESDAY

$7(3.15 - 2.98 - 3.13 + 3.18)^2 = 0.3388$

$7(2.64 - 2.98 - 3.17 + 3.18)^2 = 0.7623$

$7(3.09 - 2.98 - 3.17 + 3.18)^2 = 0.1008$

$6(3.48 - 2.98 - 3.34 + 3.18)^2 = 0.6936$

$7(3.41 - 2.98 - 3.11 + 3.18)^2 = 1.7500$

$3.6455 = \text{Subtotal } SS_{\text{Tue.}}$

WEDNESDAY

$7(3.28 - 2.93 - 3.13 + 3.18)^2 = 1.1200$

$7(2.94 - 2.93 - 3.17 + 3.18)^2 = 0.0028$

$7(2.71 - 2.93 - 3.17 + 3.18)^2 = 0.3087$

$6(3.83 - 2.93 - 3.34 + 3.18)^2 = 3.2856$

$6(3.01 - 2.93 - 3.11 + 3.18)^2 = 0.1350$

$4.8521 = \text{Subtotal } SS_{\text{Wed.}}$

THURSDAY

$6(3.50 - 3.57 - 3.13 + 3.18)^2 = 0.0024$

$6(3.33 - 3.57 - 3.17 + 3.18)^2 = 0.3174$

$6(2.95 - 3.57 - 3.17 + 3.18)^2 = 2.2326$

$6(3.50 - 3.57 - 3.34 + 3.18)^2 = 0.3174$

$6(3.43 - 3.57 - 3.11 + 3.18)^2 = 0.0294$

$2.8992 = \text{Subtotal } SS_{\text{Thur.}}$

FRIDAY

$4(2.79 - 3.36 - 3.13 + 3.18)^2 = 1.0816$

$4(2.81 - 3.36 - 3.17 + 3.18)^2 = 1.1664$

$4(2.81 - 3.36 - 3.17 + 3.18)^2 = 1.1664$

$4(3.68 - 3.36 - 3.34 + 3.18)^2 = 0.1024$

$4(3.48 - 3.36 - 3.11 + 3.18)^2 = 0.1444$

$3.6612 = \text{Subtotal } SS_{\text{Fri.}}$

GRAND TOTAL SUM OF SQUARES FOR INTERACTION DAYS BY CLASSES

Monday subtotal = 2.1794

Tuesday subtotal = 3.6455

Wednesday subtotal = 4.8521

Thursday subtotal = 2.8992

Friday subtotal = 3.6612

17.2374 = GRAND TOTAL $SS_{d \times c}$ MONDAY THROUGH FRIDAY

Table Number 5

INTERACTION: DAYS BY WEEKS

ANALYSIS OF VARIANCE—
Computation For Sums Of Squares

$SS_{d \times w}$ ACCORDING TO:

$$\sum_{i=1}^{5} \sum_{k=1}^{7} \frac{1}{j} \quad (\bar{X}_{i \cdot k} - \bar{X}_{i \cdot \cdot} - \bar{X}_{\cdot \cdot k} + \bar{\bar{X}}_{\cdot \cdot \cdot})^2$$

Number of scores in block $\Big($Mean of individual day − Mean of day over the seven weeks − Mean of all class over individual week + GM$\Big)^2$

WEEK NUMBER ONE

$5(3.71 - 3.13 - 3.36 + 3.18)^2 = 0.8000$

$5(3.30 - 3.17 - 3.36 + 3.18)^2 = 0.0125$

$5(3.14 - 3.17 - 3.36 + 3.18)^2 = \underline{0.2205}$

$1.0330 = \text{Subtotal } SS_{d \times w}$

WEEK NUMBER TWO

$5(3.95 - 3.13 - 3.92 + 3.18)^2 = 0.0320$

$5(4.10 - 3.17 - 3.92 + 3.18)^2 = 0.5445$

$3(3.67 - 3.17 - 3.92 + 3.18)^2 = 0.1728$

$5(3.98 - 3.34 - 3.92 + 3.18)^2 = \underline{0.0500}$

$0.7993 = \text{Subtotal } SS_{d \times w}$

WEEK NUMBER THREE

$4(3.96 - 3.17 - 3.78 + 3.18)^2 = 0.1444$

$5(3.60 - 3.17 - 3.78 + 3.18)^2 = 0.1445$

$5(3.77 - 3.14 - 3.78 + 3.18)^2 = \underline{0.0045}$

$0.2934 = \text{Subtotal } SS_{d \times w}$

WEEK NUMBER FOUR

$$5(3.74 - 3.17 - 3.43 + 3.18)^2 = 0.5120$$

$$5(3.23 - 3.17 - 3.43 + 3.18)^2 = 0.1805$$

$$5(3.73 - 3.34 - 3.43 + 3.18)^2 = 0.0980$$

$$5(3.01 - 3.11 - 3.43 + 3.18)^2 = 0.6125$$

$1.4033 = \text{Subtotal } SS_{d\ X\ w}$

WEEK NUMBER FIVE

$$5(1.79 - 3.13 - 2.65 + 3.18)^2 = 3.2805$$

$$5(3.26 - 3.17 - 2.65 + 3.18)^2 = 1.9220$$

$$5(2.48 - 3.17 - 2.65 + 3.18)^2 = 0.0256$$

$$5(2.61 - 3.34 - 2.65 + 3.18)^2 = 0.2000$$

$$5(3.09 - 3.11 - 2.65 + 3.18)^2 = 0.1300$$

$5.5581 = \text{Subtotal } SS_{d\ X\ w}$

WEEK NUMBER SIX

$$4(3.06 - 3.13 - 2.89 + 3.18)^2 = 0.1936$$

$$5(2.16 - 3.17 - 2.89 + 3.18)^2 = 2.5920$$

$$5(3.14 - 3.17 - 2.89 + 3.18)^2 = 0.3380$$

$$5(2.19 - 3.34 - 2.89 + 3.18)^2 = 3.6980$$

$$5(3.19 - 3.11 - 2.89 + 3.18)^2 = 0.6845$$

$7.5061 = \text{Subtotal } SS_{d\ X\ w}$

WEEK NUMBER SEVEN

$$5(3.14 - 3.13 - 2.79 + 3.18)^2 = 0.8000$$

$$5(1.66 - 3.17 - 2.79 + 3.18)^2 = 6.2720$$

$$5(2.93 - 3.17 - 2.79 + 3.18)^2 = 0.1125$$

$$5(3.06 - 3.34 - 2.79 + 3.18)^2 = 0.0121$$

$$5(3.17 - 3.11 - 2.79 + 3.18)^2 = 1.0125$$

$8.2091 = \text{Subtotal } SS_{d\ X\ w}$

GRAND TOTAL SUM OF SQUARES FOR INTERACTION DAYS BY WEEKS

Week Number One subtotal = 1.0330

Week Number Two subtotal = 0.7993

Week Number Three subtotal = 0.2934

Week Number Four subtotal = 1.4033

Week Number Five subtotal = 5.5581

Week Number Six subtotal = 7.5061

Week Number Seven subtotal = 8.2091

24.8023 = GRAND TOTAL $SS_{d \text{ X } w}$ WEEK NUMBER ONE THROUGH WEEK NUMBER SEVEN

Table Number 6

INTERACTION: CLASS(ES) BY WEEK ONE (I)

ANALYSIS OF VARIANCE— Computation For Sums Of Squares

$SS_{c \times w_1}$ ACCORDING TO:

$$\sum \text{N scores} \; (\bar{X}_{.jk} - \bar{X}_{.j.} - \bar{X}_{..k} + \bar{\bar{X}}_{...})^2$$

$$\sum \begin{pmatrix}\text{Number}\\ \text{scores}\end{pmatrix} \times \left[\begin{pmatrix}\text{Vary the classes}\\ \text{I through 5 over}\\ \text{the weeks I}\\ \text{through 7}\end{pmatrix} - \begin{pmatrix}\text{Mean for}\\ \text{classes}\end{pmatrix} - \begin{pmatrix}\text{Mean}\\ \text{for}\\ \text{weeks}\end{pmatrix} + \begin{pmatrix}\text{GM}\end{pmatrix} \right]^2$$

CLASS	WEEK NUMBER ONE
I	$3(9.67 / 3 - 3.06 - 3.36 + 3.18)^2 = 0.0012$
II	$3(8.41 / 3 - 2.98 - 3.36 + 3.18)^2 = 0.3888$
III	$3(9.50 / 3 - 2.93 - 3.36 + 3.18)^2 = 0.0108$
V	$3(12.53 / 3 - 3.57 - 3.36 + 3.18)^2 = 0.5547$
VI	$3(10.63 / 3 - 3.36 - 3.36 + 3.18)^2 = 0.0000$

0.9555 = Subtotal $SS_{c \times w_1}$

INTERACTION: CLASS(ES) BY WEEK TWO (2)

CLASS	WEEK NUMBER TWO
I	$4(15.85 / 4 - 3.06 - 3.92 + 3.18)^2 = 0.1024$
II	$4(15.57 / 4 - 2.98 - 3.92 + 3.18)^2 = 0.1156$
III	$4(14.99 / 4 - 2.93 - 3.92 + 3.18)^2 = 0.0100$
V	$3(12.84 / 3 - 3.57 - 3.92 + 3.18)^2 = 0.0027$
VI	$3(11.94 / 3 - 3.36 - 3.92 + 3.18)^2 = 0.0432$
	0.2749 = Subtotal $SS_{c\ X\ w_2}$

INTERACTION: CLASS(ES) BY WEEK THREE (3)

CLASS	WEEK NUMBER THREE
I	$3(11.25 / 3 - 3.06 - 3.78 + 3.18)^2 = 0.0243$
II	$3(10.72 / 3 - 2.98 - 3.78 + 3.18)^2 = 0.0003$
III	$3(10.49 / 3 - 2.93 - 3.78 + 3.18)^2 = 0.0147$
V	$2(08.66 / 2 - 3.57 - 3.78 + 3.18)^2 = 0.0512$
VI	$3(11.61 / 3 - 3.36 - 3.78 + 3.18)^2 = 0.0243$
	0.1148 = Subtotal $SS_{c\ X\ w_3}$

INTERACTION: CLASS(ES) BY WEEK FOUR (4)

CLASS	WEEK NUMBER FOUR
I	$4(14.72 / 4 - 3.06 - 3.43 + 3.18)^2 = 0.5476$
II	$4(12.41 / 4 - 2.98 - 3.43 + 3.18)^2 = 0.0676$
III	$4(13.76 / 4 - 2.93 - 3.43 + 3.18)^2 = 0.0784$
V	$4(15.17 / 4 - 3.57 - 3.43 + 3.18)^2 = 0.0036$
VI	$4(11.96 / 4 - 3.36 - 3.43 + 3.18)^2 = 0.5476$
	1.2448 = Subtotal $SS_{c\ X\ w_4}$

INTERACTION: CLASS(ES) BY WEEK FIVE (5)

CLASS	WEEK NUMBER FIVE
I	$5(11.15 / 5 - 3.06 - 2.65 + 3.18)^2 = 0.4500$
II	$5(12.57 / 5 - 2.98 - 2.65 + 3.18)^2 = 0.0180$
III	$5(11.99 / 5 - 2.93 - 2.65 + 3.18)^2 = 0.0000$
V	$5(15.42 / 5 - 3.57 - 2.65 + 3.18)^2 = 0.0080$
VI	$5(15.04 / 5 - 3.36 - 2.65 + 3.18)^2 = 0.1620$

$0.6380 = \text{Subtotal } SS_c \times w_5$

INTERACTION: CLASS(ES) BY WEEK SIX (6)

CLASS	WEEK NUMBER SIX
I	$5(14.41 / 5 - 3.06 - 2.89 + 3.18)^2 = 0.0605$
II	$5(13.70 / 5 - 2.98 - 2.89 + 3.18)^2 = 0.0180$
III	$5(12.20 / 5 - 2.93 - 2.89 + 3.18)^2 = 0.2000$
V	$4(12.24 / 4 - 3.57 - 2.89 + 3.18)^2 = 0.1936$
VI	$5(16.75 / 5 - 3.36 - 2.89 + 3.18)^2 = 0.3920$

$0.8461 = \text{Subtotal } SS_c \times w_6$

INTERACTION: CLASS(ES) BY WEEK SEVEN (7)

CLASS	WEEK NUMBER SEVEN
I	$5(13.27 / 5 - 2.06 - 2.79 + 3.18)^2 = 0.0020$
II	$5(12.76 / 5 - 2.98 - 2.79 + 3.18)^2 = 0.0080$
III	$5(12.61 / 5 - 2.93 - 2.79 + 3.18)^2 = 0.0020$
V	$5(16.25 / 5 - 3.57 - 2.79 + 3.18)^2 = 0.0245$
VI	$5(14.90 / 5 - 3.36 - 2.79 + 3.18)^2 = 0.0005$

$0.0370 = \text{Subtotal } SS_c \times w_7$

Chart Number 7

SUMMARY CHART: SS CLASS BY WEEKS
INTERACTION FOR $c \times w_{k=1-7}$

WEEK

(1) $c \times w_1$ Interaction SS = 0.9555

(2) $c \times w_2$ Interaction SS = 0.2749

(3) $c \times w_3$ Interaction SS = 0.1148

(4) $c \times w_4$ Interaction SS = 1.2448

(5) $c \times w_5$ Interaction SS = 0.6380

(6) $c \times w_6$ Interaction SS = 0.8461

(7) $c \times w_7$ Interaction SS = 0.0370

4.1111 = GRAND TOTAL $SS_{c \times w_{k=1-7}}$

Table Number 8

MAIN EFFECT: DAYS
Calculation For F-Ratio

F-ratios for $MS_d = 0.2332$ at 4 over (62 + 24) degrees of freedom.

$$F_d = \frac{MS_d}{MS_{res} + MS_{c \times w}}$$

$$= \frac{0.2332}{0.3308 + 0.2569}$$

CALCULATED F_d VALUE = ...0.3968

	RESULTS
Table F Value @ 4/80 D.F. for 1% confidence level = ...3.56	$> F_d$
Table F Value @ 4/80 D.F. for 5% confidence level = ...2.48	$> F_d$
Table F Value @ 4/100 D.F. for 1% confidence level = ...3.51	$> F_d$
Table F Value @ 4/100 D.F. for 5% confidence level = ...2.46	$> F_d$

ACCORDING TO RATIO COMPARISON:

No significance revelaed by analysis and comparison of these obtained data at the confidence levels indicated. The investigators failed to reject H_0 at levels shown above.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value $>$ Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied—unable to reject H_0.

Table Value $<$ Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied—ability to reject H_0 is indicated.

Table Number 9

MAIN EFFECT: CLASSES
Calculation For F-Ratio

F-ratios for $MS_c = 1.8837$ at 4 over $(62 + 24)$ degrees of freedom.

$$F_c = \frac{MS_c}{MS_{res} + MS_{d \times w}}$$

$$= \frac{1.8837}{0.3308 + 1.0334}$$

CALCULATED F_c VALUE = ...1.3809

RESULTS

Table F Value @ 4/80 D.F. for 1% confidence level = ...3.56 $> F_c$

Table F Value @ 4/80 D.F. for 5% confidence level = ...2.48 $> F_c$

Table F Value @ 4/100 D.F. for 1% confidence level = ...3.51 $> F_c$

Table F Value @ 4/100 D.F. for 5% confidence level = ...2.46 $> F_c$

ACCORDING TO RATIO COMPARISON:

No significance revealed by analysis and comparison of these obtained data at the confidence levels indicated. The investigators failed to reject H_0 at levels shown above.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value > Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied—unable to reject H_0.

Table Value < Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied—ability to reject H_0 is indicated.

Table Number 10

MAIN EFFECT: WEEKS
Calculation For F-Ratio

F-ratios for MS_w = 1.3857 at 6 over (62 + 22) degrees of freedom.

$$F_w = \frac{MS_w}{MS_{res} + MS_{c \times d}}$$

$$= \frac{1.3857}{0.3308 + 0.7186}$$

CALCULATED F_w VALUE = ...1.3205

		RESULTS
Table F Value @ 6/80 D.F. for 1% confidence level =	3.04	$> F_w$
Table F Value @ 6/80 D.F. for 5% confidence level =	2.21	$> F_w$
Table F Value @ 6/100 D.F. for 1% confidence level =	2.99	$> F_w$
Table F Value @ 6/100 D.F. for 5% confidence level =	2.19	$> F_w$

ACCORDING TO RATIO COMPARISON:

No significance revealed by analysis and comparison of these obtained data at the confidence levels indicated. The investigators failed to reject H_0 at levels shown above.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value > Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied—unable to reject H_0.

Table Value < Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied—ability to reject H_0 is indicated.

Table Number 11

INTERACTION: DAYS VS. WEEKS
Calculation For F-Ratio

F-ratios for $MS_{d X w} = 1.0334$ at 24 over 62 degrees of freedom

$$F_{d X w} = \frac{MS_{d X w}}{MS_{res}}$$

$$= \frac{1.0334}{0.3308}$$

CALCULATED $F_{d X w}$ VALUE = ...3.1239

		RESULTS
Table F Value @ 24/60 D.F. for 1% confidence level =	..........................2.12	$< F_{d X w}$
Table F Value @ 24/60 D.F. for 5% confidence level =	..........................1.70	$< F_{d X w}$
Table F Value @ 24/65 D.F. for 1% confidence level =	..........................2.09	$< F_{d X w}$
Table F Value @ 24/65 D.F. for 5% confidence level =	..........................1.68	$< F_{d X w}$

ACCORDING TO RATIO COMPARISON:

The analysis and comparison to which these data were subjected indicate a significant variance related to interaction of the considered effects at the above levels--the investigators were able, therefore, to reject H_0 at all confidence levels indicated above.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value $>$ Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied--unable to reject H_0.

Table Value $<$ Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied--ability to reject H_0 is indicated.

Table Number 12

INTERACTION: DAYS VS. CLASSES
Calculation For F-Ratio

F-ratios for $MS_{d \times c}$ = 0.7186 at 16 over 62 degrees of freedom.

$$F_{c \times d} = \frac{MS_{d \times c}}{MS_{res}}$$

$$= \frac{0.7186}{0.3308}$$

CALCULATED $F_{d \times c}$ VALUE = ...2.1720

RESULTS

Table F Value @ 16/60 D.F. for 1% confidence level = ...2.32 $> F_{d \times c}$

Table F Value @ 16/60 D.F. for 5% confidence level = ...1.81 $< F_{d \times c}$

Table F Value @ 16/65 D.F. for 1% confidence level = ...2.30 $> F_{d \times c}$

Table F Value @ 16/65 D.F. for 5% confidence level = ...1.80 $< F_{d \times c}$

ACCORDING TO RATIO COMPARISON:

1. The above determinations reveal some degree of interaction related to days vs. classes that appear to contribute to a significant variation at the .05 level for these data analyzed.*
2. No significance revealed for these data at the 1% level.**

*Rejected H_0 at indicated levels of confidence for these data.

**Failed to reject H_0 at levels indicated for these data.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value > Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied—unable to reject H_0.

Table Value < Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied—ability to reject H_0 is indicated.

Table Number 13

INTERACTION: CLASS VS. WEEKS
Calculation For F-Ratio

F-ratios for $MX_{c X w} = 0.2569$ at 24 over 62 degrees of freedom.

$$F_{c X w} = \frac{MS_{c X w}}{MS_{res}}$$

$$= \frac{0.2569}{0.3308}$$

CALCULATED $F_{c X w}$ VALUE = ...0.7766

	RESULTS
Table F Value @ 24/60 D.F. for 1% confidence level = ...2.12	$> F_{c X w}$
Table F Value @ 24/60 D.F. for 5% confidence level = ...1.70	$> F_{c X w}$
Table F Value @ 24/65 D.F. for 1% confidence level = ...2.09	$> F_{c X w}$
Table F Value @ 24/65 D.F. for 5% confidence level = ...1.68	$> F_{c X w}$

ACCORDING TO RATIO COMPARISON:

No significance revealed by analysis and comparison of these obtained data at the confidence levels indicated. The investigators failed to reject H_0 at levels shown above.

KEY FOR F DETERMINATION AND ANALYSIS:

Table Value $>$ Computed Value = No significant variation in data obtained under the conditions considered according to determinations and analysis applied—unable to reject H_0.

Table Value $<$ Computed Value = There is a significant variation in data obtained under the conditions considered according to determinations and analysis applied—ability to reject H_0 is indicated.

Chart Number 14

ANOVA: DETAILED SUMMARY CHART FOR INTERACTIONS AS SOURCES OF ERRORS

SOURCE OF INTERACTION	MS	DEGREES OF FREEDOM RATIO AS COMPUTED	D.F.	RESULT OF COMPARISON WITH FISHER'S F TABLE*		
	SS/DF	"Between/within"		Calculation F Value	1%	5%
DAYS BY WEEKS	1.0334	24/62	24/60	3.1239	(+)	(+)
			24/65		(+)	(+)
CLASSES BY DAYS	0.7186	16/62	16/60	2.1720	(-)	(+)
			16/65		(-)	(+)
CLASSES BY WEEKS	0.2569	24/62	24/60	0.7766	(-)	(-)
			24/65		(-)	(-)

* (+) = Significance IS indicated at the confidence level(s) considered.

(-) = Significance IS NOT indicated at the confidence level(s) considered.

Chart Number 15

ANOVA: SUMMARY CHART FOR MAIN EFFECTS AS SOURCES OF ERRORS

MAIN EFFECT SOURCE	MS	DEGREES OF FREEDOM RATIO AS COMPUTED	D.F.	RESULT OF COMPARISON WITH FISHER'S F TABLE*		
	SS/DF	"Between/ within"		Calculation F Value	1%	5%
DAYS	0.2332	4/86	4/80	0.3968	(-)	(-)
			4/100		(-)	(-)
CLASSES	1.8837	4/86	4/80	1.3809	(-)	(-)
			4/100		(-)	(-)
WEEKS	1.3857	6/84	6/80	1.3205	(-)	(-)
			6/80		(-)	(-)

* (+) = Significance IS indicated at the confidence level(s) considered.

(-) = Significance IS NOT indicated at the confidence level(s) considered.

RESIDUAL VARIANCE

ACCORDING TO:

$$\text{RESIDUAL}_{\text{var}} = \text{TOTAL}_{\text{var}} - \Big[\text{DAY}_{\text{var}} + \text{CLASSES}_{\text{var}} + \text{WEEKS}_{\text{var}} + (\text{DAYS X WEEKS})_{\text{var}} + (\text{DAYS X CLASSES})_{\text{var}} + (\text{CLASSES X WEEKS})_{\text{var}}\Big]$$

$$\text{RESIDUAL}_{\text{var}} = 83.4441 - (0.9327 + 7.5350 + 8.3144 + 24.8023 + 4.1111 + 17.2374)$$

$$\text{RESIDUAL}_{\text{var}} = 20.5122$$

Chart Number 16

SUMMARY TABLE FOR ANALYSIS OF VARIANCE OF ERROR SOURCES

ERROR SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	COMPUTED SIGNIFICANCE RATIO VALUE	D.F. RATIO FOR COMPUTATIONAL PURPOSES[1]	RESULTS OF FISHER'S SIGNIFICANCE RATIO VALUE[2]
MAIN EFFECTS	SS	-DF	=MS	F		
DAYS: (D) i=1,2,3, 4,5	0.9327	4	0.2332	0.3968	4:86	(>)*
CLASSES: (C) j= 1,2,3,4,5	7.5350	4	1.8837	1.3809	4:86	(>)*
WEEKS: (W) k=1, 2,3,4, 5,6,7	8.3144	6	1.3857	1.3205	6:84	(>)*
INTER-ACTIONS	SS	-DF	=MS	F		
D X W i= k= 1-5 1-7	24.8023	24	1.0334	3.1239	24:62	(<)**
D X C i= j= 1-5 1-5	17.2374	16	0.7120	2.1720	24:62	(<)**
C X W i= j= 1-5 1-5	4.1111	24	0.2569	0.7766	16:62	(>)*
RES_{var}	20.5112	62	0.3308			
$TOTAL_{var}$	83.4441	140	0.5960			

*Based on the above analysis of these obtained data for this study, there is no indication of a significant variation at the confidence level(s) indicated.

**Based on the above analysis of these obtained data for this study, there is indication of a significant variation at the confidence level(s) indicated.

---

[1] As defined and used in this supplementary report.

[2] Herbert Arkin and Raymond R. Colton, Tables For Statisticians, New York: Barnes & Noble, Inc., 1962, pp. 117-21. Comparison based on Snedecor's modified Fisher's Table of F Values.

CONCLUSIONS

Based on the evaluation as contained in this report the writer concludes that the sources of between groups variance were traceable to an indicated significance at the point of interaction of:

1. DAYS with WEEKS.
2. DAYS with CLASSES.

The analysis not only allowed the experimenters to detect significant performance difference related to differential treatment at the levels indicated but also gave some information as to which areas and/or combination of areas were more involved in these detected variations.

Since the combined variables of: (1) days with weeks and (2) days with classes accounted for a greater portion of the variance than was obtained by adding together the individual estimations of variance for each variable under consideration, the interaction was determined to be significant and confirmed by F table comparisons.

No other variable relationships analyzed in this study revealed any indication of significance at the confidence levels considered.

INDICATIONS FOR THE FUTURE

According to Gage:

> "The second erroneous notion seems to be that one experiment can and ought to shed light on only one hypothesis..."

Although few if any solutions may be found within the limitations of this study, it is felt that MANY related areas for potential investigation have been presented and/or suggested. This investigator plans more precise and more controlled effort along these lines in the very near future and hopes that others will also find some interest in the suggested areas of investigation.

ANOVA - DAY(S) X WEEK(S)
($\bar{c}$ Nested Classes)

Week	Period Class	A Monday		B Tuesday		C Wednesday		D Thursday		E Friday		Σ's			MEANS
	I	3.60	3/25/68	2.69	3/26/68	3.38	3/27/68		3/28/68		3/29/68	ΣI	9.67		
	II	3.59		3.00		1.82						ΣII	8.41		
1.	III	3.60	Σ = 18.56	3.42	Σ = 16.50	2.48	Σ = 15.68					ΣIII	9.50		$\bar{x}$ =16.91
	V	3.87	$\bar{x}$ = 3.71	4.29	$\bar{x}$ = 3.30	4.37	$\bar{x}$ = 3.14					ΣV	12.53		
	VI	3.90		3.10		3.63						ΣVI	10.63	Σ =50.74	$\bar{\bar{x}}$ = 3.36
	I	3.95	4/1/68	4.40	4/2/68	3.43	4/3/68	4.07	4/4/68		4/5/68	ΣI	15.85		
	II	4.32		3.32		3.82		4.11				ΣII	15.57		
2.	III	3.25	Σ = 19.77	4.03	Σ = 20.52	3.75	Σ = 11.00	3.96	Σ = 19.90			ΣIII	14.99		$\bar{x}$ =17.80
	V	4.30	$\bar{x}$ = 3.95	4.54	$\bar{x}$ = 4.10	----	$\bar{x}$ = 3.67	4.00	$\bar{x}$ = 3.98			ΣV	12.84		
	VI	3.95		4.23		----		3.76				ΣVI	11.94	Σ =71.19	$\bar{\bar{x}}$ = 3.92
	I		4/8/68	4.11	4/9/68	3.62	4/10/68	3.52	4/11/68		4/12/68	ΣI	11.25		
	II			3.86		2.86		4.00				ΣII	10.72		
3.	III			3.74	Σ = 15.86	3.06	Σ = 18.02	3.69	Σ = 18.85			ΣIII	10.49		$\bar{x}$ =17.58
	V			----	$\bar{x}$ = 3.96	4.62	$\bar{x}$ = 3.60	4.04	$\bar{x}$ = 3.77			ΣV	8.66		
	VI			4.15		3.86		3.60				ΣVI	11.61	Σ =52.73	$\bar{\bar{x}}$ = 3.78
	I		4/15/68	3.80	4/16/68	4.00	4/17/68	3.85	4/18/68	3.07	4/19/68	ΣI	14.72		
	II			3.00		3.78		3.06		2.57	Σ = 15.03	ΣII	12.41		
4.	III			3.50	Σ = 18.69	3.39	Σ = 16.95	3.80	Σ = 18.65	2.57		ΣIII	13.26		$\bar{x}$ =17.13
	V			4.08	$\bar{x}$ = 3.74	3.17	$\bar{x}$ = 3.23	4.14	$\bar{x}$ = 3.73	3.78	$\bar{x}$ = 3.01	ΣV	15.17		
	VI			4.31		1.81		3.80		3.04		ΣVI	12.96	Σ =68.52	$\bar{\bar{x}}$ = 3.43
	I	1.33	4/22/68	2.40	4/23/68	1.82	4/24/68	3.12	4/25/68	2.48	4/26/68	ΣI	11.15		
	II	2.30		2.57		2.17		2.57		2.96		ΣII	12.57		
5.	III	1.72	Σ = 8.97	3.14	Σ = 16.29	1.91	Σ = 12.41	2.15	Σ = 13.04	3.07	Σ = 15.46	ΣIII	11.99		$\bar{x}$ =13.23
	V	1.58	$\bar{x}$ = 1.79	4.13	$\bar{x}$ = 3.26	3.65	$\bar{x}$ = 2.48	2.20	$\bar{x}$ = 2.61	3.86	$\bar{x}$ = 3.09	ΣV	15.42		
	VI	2.04		4.05		2.86		3.00		3.09		ΣVI	15.04	Σ = 66.17	$\bar{\bar{x}}$ = 2.65
	I	1.92	4/29/68	2.76	4/30/68	3.70	5/1/68	2.84	5/2/68	3.19	5/3/68	ΣI	14.41		
	II	2.38		1.65		3.42		3.60		2.65		ΣII	13.70		
6.	III	3.23	Σ = 12.26	1.84	Σ = 10.81	2.23	Σ = 15.70	2.17	Σ = 14.56	2.73	Σ = 15.97	ΣIII	12.20		$\bar{x}$ =13.86
	V	----	$\bar{x}$ = 3.06	2.04	$\bar{x}$ = 2.16	3.50	$\bar{x}$ = 3.14	2.90	$\bar{x}$ = 2.91	3.80	$\bar{x}$ = 3.19	ΣV	12.24		
	VI	4.73		2.52		2.85		3.05		3.60		ΣVI	16.75	Σ =69.30	$\bar{\bar{x}}$ = 2.89
	I	2.26	5/6/68	1.92	5/7/68	3.04	5/8/68	3.61	5/9/68	2.44	5/10/68	ΣI	13.27		
	II	3.27		1.07		2.71		2.64		3.07		ΣII	12.76		
7.	III	3.61	Σ = 15.70	2.00	Σ = 8.29	2.19	Σ = 14.67	1.96	Σ = 15.29	2.86	Σ = 15.85	ΣIII	12.62		$\bar{x}$ =13.96
	V	3.76	$\bar{x}$ = 3.14	1.80	$\bar{x}$ = 1.66	3.69	$\bar{x}$ = 2.93	3.70	$\bar{x}$ = 3.06	3.30	$\bar{x}$ = 3.17	ΣV	16.25		
	VI	2.80		1.50		3.04		3.38		4.18		ΣVI	14.90	Σ =69.80	$\bar{\bar{x}}$ = 2.79

$\bar{\bar{\Sigma}}_m = 75.26$	$\bar{\bar{\Sigma}}_t = 106.96$	$\bar{\bar{\Sigma}}_w = 103.63$	$\bar{\bar{\Sigma}}_t = 100.29$	$\bar{\bar{\Sigma}}_f = 62.31$	$\bar{\bar{\Sigma}} = 448.45$ $\bar{N} = 141$
$\Sigma X = 15.65$	$\Sigma X = 22.18$	$\Sigma X = 22.19$	$\Sigma X = 20.06$	$\Sigma X = 12.46$	
$(\Sigma X)^2 = 244.92$	$(\Sigma X)^2 = 491.95$	$(\Sigma X)^2 = 492.40$	$(\Sigma X)^2 = 402.40$	$(\Sigma X)^2 = 155.25$	GRAND MEAN = 3.18
$\bar{\bar{X}}_m = 3.13$	$\bar{\bar{X}}_t = 3.17$	$\bar{\bar{X}}_w = 3.17$	$\bar{\bar{X}}_t = 3.34$	$\bar{\bar{X}}_f = 3.11$	

Chart Constructed By: Donna Hughes Joyner

ANOVA: CLASS(S) X DAY(S)
(C WEEKS NESTED)

		I		II		III		V		VI		$\sum\sum$	$\sum \bar{X}$
Mon	Week											$\sum\sum = 75.26$	$\sum \bar{X} = 15.05$
	1	3.60		3.59		3.60		3.87		3.90			
	2	3.95		4.32		3.25		4.30		3.95			
	3	——	$\sum = 13.06$	——	$\sum = 22.08$	——	$\sum = 15.41$	——	$\sum = 13.51$	——	$\sum = 17.42$		
	4	——		——		——		——		——			
	5	1.33		2.30		1.72		1.58		2.04			
	6	1.92	$\bar{X} = 2.61$	2.38	$\bar{X} = 3.17$	3.23	$\bar{X} = 3.08$	——	$\bar{X} = 3.38$	4.73	$\bar{X} = 3.48$		
	7	2.26		3.27		3.61		3.76		2.80			
Tue	Week											$\sum\sum = 106.96$	$\sum \bar{X} = 23.34$
	1	2.69		3.00		3.42		4.29		3.10			
	2	4.40	$\sum = 22.08$	3.32	$\sum = 18.47$	4.03	$\sum = 21.67$	4.54	$\sum = 20.88$	4.23	$\sum = 23.86$		
	3	4.11		3.86		3.74		——		4.15			
	4	3.80	$\bar{X} = 3.15$	3.00	$\bar{X} = 2.64$	3.50		4.08	$\bar{X} = 3.48$	4.31	$\bar{X} = 3.41$		
	5	2.40		2.57		3.14	$\bar{X} = 3.09$	4.13		4.05			
	6	2.76		1.65		1.84		2.04		2.52			
	7	1.92		1.07		2.00		1.80		1.50			
Wed	Week											$\sum\sum = 103.63$	$\sum \bar{X} = 20.73$
	1	3.38		1.82		2.48		4.37		3.63			
	2	3.43	$\sum = 22.99$	3.82	$\sum = 20.58$	3.75	$\sum = 19.01$	——	$\sum = 23.00$	——	$\sum = 18.05$		
	3	3.62		2.86		3.06		4.62		3.86			
	4	4.00		3.78		3.39		3.17		1.81			
	5	1.82	$\bar{X} = 3.28$	2.17	$\bar{X} = 2.94$	1.91	$\bar{X} = 2.71$	3.65	$\bar{X} = 3.83$	2.86	$\bar{X} = 3.01$		
	6	3.70		3.42		2.23		3.50		2.85			
	7	3.04		2.71		2.19		3.69		3.04			
Thu	Week											$\sum\sum = 100.29$	$\sum \bar{X} = 20.06$
	1	——		——		——		——		——			
	2	4.07	$\sum = 21.01$	4.11	$\sum = 19.98$	3.96	$\sum = 17.73$	4.00	$\sum = 20.98$	3.76	$\sum = 20.59$		
	3	3.52		4.00		3.69		4.04		3.60			
	4	3.82		3.06		3.80		4.14		3.80			
	5	3.12	$\bar{X} = 3.50$	2.57	$\bar{X} = 3.33$	2.15	$\bar{X} = 2.95$	2.20	$\bar{X} = 3.50$	3.00	$\bar{X} = 3.43$		
	6	2.84		3.60		2.17		2.90		3.05			
	7	3.61		2.64		1.96		3.70		3.38			
Fri	Week											$\sum\sum = 62.31$	$\sum \bar{X} = 12.46$
	1	——		——		——		——		——			
	2	——	$\sum = 11.18$	——	$\sum = 11.25$	——	$\sum = 11.23$	——	$\sum = 14.74$	——	$\sum = 13.91$		
	3	——		——		——		——		——			
	4	3.07		2.57		2.57		3.78		3.04			
	5	2.48	$\bar{X} = 2.79$	2.96	$\bar{X} = 2.81$	3.07	$\bar{X} = 2.81$	3.86	$\bar{X} = 3.68$	3.09	$\bar{X} = 3.48$		
	6	3.19		2.65		2.73		3.80		3.60			
	7	2.44		3.07		2.86		3.30		4.18			
		$\sum = 90.32$		$\sum = 86.14$		$\sum = 85.05$		$\sum = 93.11$		$\sum = 93.83$			
		$\sum \bar{X} = 15.33$		$\sum \bar{X} = 14.89$		$\sum \bar{X} = 14.64$		$\sum \bar{X} = 17.87$		$\sum \bar{X} = 16.81$			
		$\bar{\bar{X}} = 3.06$		$\bar{\bar{X}} = 2.98$		$\bar{\bar{X}} = 2.93$		$\bar{\bar{X}} = 3.57$		$\bar{\bar{X}} = 3.36$			

$\sum\sum = 448.45$

$\bar{N} = 141$ GM = $\underline{3.18}$

Charting by: Donna H. Joyner

BIBLIOGRAPHY

Arkin, Herbert and Colton, Raymond R. Statistical Methods. New York: Barnes & Noble, Inc. 1967.

Arkin, Herbert and Colton, Raymond R. Tables for Statisticians. New York: Barnes & Noble, Inc. 1962.

Baker, Robert L. and Schutz, Richard E. Analyzing Variability. Inglewood, California: Southwest Regional Laboratory for Educational Research and Development. 1967.

Campbell, Donald T. and Stanley, Julian C. Experimental And Quasi—Experimental Designs For Research. Chicago: Rand McNally & Company. 1966.

Cohen, Dorothy H. and Stern, Virginia. Observing and Recording the Behavior of Young Children. Teachers College, Columbia University: Teachers College Press. 1965.

Courtney, E. Wayne. Applied Research In Education. Totowa, New Jersey: Littlefield, Adams & Co. 1965.

Downie, N. M. Fundamentals Of Measurement. New York: Oxford University Press. 1967.

Edwards, Allen L. Statistical Methods. New York: Holt, Rinehart and Winston, Inc. 1967.

Gadske, R. E. "A Comparison of two Methods of Teaching First Year High School Algebra." Sch. Sci. Math. 1933.

Gage, N. L. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally & Company. 1967.

Garrett, Henry E. Testing for Teachers. New York: American Book Company, Inc. 1961.

Herink, Reinhard. College Level Statistics. New York: Monarch Monarch Press, Inc. 1965

Herink, Reinhard and Feldman, Harold. Statistics. New York: Monarch Press, Inc. 1965.

Josselyn, Irene M. Psychosocial Development of Children. New York: Family Service Association of America. 1954.

Massialas, Byron G. and Smith, Frederick R. "Quality Research — A Goal For Every Teacher," Applied Research in Education. Totowa, New Jersey: Littlefield, Adams & Co. 1965.

McGuigan, F. J. Experimental Psychology. Englewood Cliffs, N.J.: Prentice-Hall, Inc. 1965.

Moroney, M. J. *Facts from Figures*. Baltimore: Pelican Books. 1956

Morse, William C. and Wingo, G. Max. *Psychology and Teaching*. Chicago: Scott, Foresman and Company. 1962

Mosteller, Frederick, Rourke, Robert E. K., and Thomas, George B. *Probability and Statistics*. London: Addison-Wesley Publishing Company, Inc. 1961.

Pintner, Rudolf, *et al*. *Educational Psychology*. New York: Barnes & Noble, Inc. 1965.

Scheffé, Henry. *The Analysis of Variance*. New York: John Wiley and Sons, Inc. 1959.

Snedecor, George W. *Statistical Methods*. Ames, Iowa: Iowa State College Press. 1946.

Speigel, Murray R. *Theory and Problems of Statistics*. New York: Schaum Publishing Co. 1961.

Wood, Dorothy Adkins. *Test Construction*. Columbus, Ohio: Charles E. Merrill Books, Incorporated. 1961.

RESEARCH TRAINING

FALL SEMESTER, 1967

Course Outline

Text Book: Downie, N. M. Fundamentals of Measurement. New York: Oxford University Press.

Part I Tests & Testing

1. Introduction
   - a. Philosophical and Psychological Principles of Measurement and Evaluation

   - b. Basic Statistical Tools and vocabulary

2. Scores, Norms and Evaluations

3. Reliability and Validity

Part II Defining Objectives

1. Science and Scientific Approach
2. Problems and Hyphotheses
3. Constructs and Variables
4. Planning Measurement and Evaluation

Part III Instruments for Measurement

1. Methods of Observation and Data Collection
2. Constructing Instruments
3. Standardized Forms
   - a. Achievement
   - b. Other Behaviorial Measures
4. Factors in the Selection of Tests
5. Measurement Ethics

Part IV Preparing for the Research Project

SUGGESTED READINGS

Part I Tests & Testing

Downie - Chapters 1, 2, 3, & 4
Kerlinger - Chapters 5, 6, 7, 8, 23, 24, & 25
Payne & McMorris - Chapters 1, 2, 3, & 4
Lien - Chapter 1
Schutz & Baker - (Mimeographed) "The Experimental Analysis of Behavior"
Garrett - Chapters 1, 2, 5, & 13
Gerlach et al - "Relative VS. Absolute Criteria"

Part II Defining Objectives

Gage - Chapter 2
Bloom's Taxonomy
Lien - Chapter 2
Kerlinger - Chapters 1, 2, & 3
Gerlach et al - "Defining Instructional Specifications"
"Describing Educational Outcomes"
"Constructing Statements of Outcomes"

Part III Instruments for Measurement

Mental Measurements Yearbook - Pages 1-1379
Kerlinger - Chapters 26, 27, 28, 29, 30, 31, 32, & 33
Webb et al - Entire Book
ETS - Test Construction Kit
Lien - Chapters 4, 5, & 6
Downie - Chapters 6-18
Payne & McMorris - Chapters 5 & 6
Conrad - Sociology of Education, Spring 1967

Part IV Preparing for the Research Project

Gage - Chapter 1
Kerlinger - Chapters 15, 16, 17, 18, & 19

BIBLIOGRAPHY
RESEARCH TRAINING COURSE # 4511
FALL SEMESTER 1967

Conrad, Herbert S. "Clearance of Questionnaires with Respect to 'Invasion of Privacy' Public Sensitivities, Ethical Standards, Etc.: Principles and Viewpoint in the Bureau of Research, U. S. Office of Education" *Sociology of Education*. Vol. 40, No. 2, Spring 1967, pp. 170-175.

Downie, N. M. *Fundamentals of Measurement: Techniques and Practices*. 2d ed. New York, London and Toronto: Oxford University Press, 1967.

Garrett, Henry E. Ph. D. *Statistics in Psychology and Education*. 6th ed. New York: David McKay Company, Inc., 1966.

Gerlach, Vernon S., and Sullivan, Howard J. *Constructing Statements of Outcomes*. Southwest Regional Laboratory for Educational Research and Development, 1967.

Gerlach, Vernon S., Baker, Robert L., Schutz, Richard E., and Sullivan, Howard J. *Defining Instructional Specifications*. Southwest Regional Laboratory for Educational Research and Development, 1967.

Gerlach, Vernon S., *Describing Educational Outcomes*. Southwest Regional Laboratory for Educational Research and Development. 1967.

Gerlach, Vernon S., Schutz, Richard E., and Baker, Robert L. *Relative VS. Absolute Criteria*. Southwest Regional Laboratory for Educational Research and Development, 1967.

Handbook of Research on Teaching. Edited by N. L. Gage. A Project of The American Educational Research Association, A Department of the National Education Association. Chicago: Rand McNally & Company, 1965.

Kerlinger, Fred N. *Foundations of Behavioral Research: Educational and Psychological Inquiry*. New York: Holt, Rinehart and Winston, Inc. 1966.

Lien, Arnold J. *Measurement and Evaluation of Learning: A Handbook for Teachers*. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1967.

Payne, David A., and McMorris, Robert F. *Educational and Psychological Measurement. Contributions to Theory and Practice*. Waltham, Massachusetts: Blaisdell Publishing Company, 1967.

Schutz, Richard E., and Baker, Robert L. *"The Experimental Analysis of Behavior" in Educational Research*. Southwest Regional Laboratory för Educational Research and Development.

Webb, Eugene J., Campbell, Donald T., Schwartz, Richard D., and Sechrest, Lee. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: Rand McNally & Company, 1966.